Pointwise stabilization of discrete-time matrix-valued stationary Markov chains[☆]

Xiongping Dai[a], Yu Huang[b], Mingqing Xiao[c]

[a]*Department of Mathematics, Nanjing University, Nanjing 210093, People's Republic of China*
[b]*Department of Mathematics, Zhongshan (Sun Yat-Sen) University, Guangzhou 510275, People's Republic of China*
[c]*Department of Mathematics, Southern Illinois University, Carbondale, IL 62901-4408, USA*

Abstract

Let $(\Omega, \mathscr{F}, \mathbb{P})$ be a probability space and $S = \{S_1, \ldots, S_K\}$ a discrete-topological space that consists of K real d-by-d matrices, where K and d both ≥ 2. In this paper, we study the pointwise stabilizability of a discrete-time, time-homogeneous, stationary (p, P)-Markovian jump linear system $\Xi = (\xi_n)_{n=1}^{+\infty}$ where $\xi_n \colon \Omega \to S$. Precisely, Ξ is called "pointwise convergent", if to any initial state $x_0 \in \mathbb{R}^{1 \times d}$, there corresponds a measurable set $\Omega_{x_0} \subset \Omega$ with $\mathbb{P}(\Omega_{x_0}) > 0$ such that

$$x_0 \prod\nolimits_{\ell=1}^{n} \xi_\ell(\omega) \to \mathbf{0}_{1 \times d} \quad \text{as } n \to +\infty, \qquad \forall \omega \in \Omega_{x_0};$$

and Ξ is said to be "pointwise exponentially convergent", if to any initial state $x_0 \in \mathbb{R}^{1 \times d}$, there corresponds a measurable set $\Omega'_{x_0} \subset \Omega$ with $\mathbb{P}(\Omega'_{x_0}) > 0$ such that

$$x_0 \prod\nolimits_{\ell=1}^{n} \xi_\ell(\omega) \xrightarrow{\text{exponentially fast}} \mathbf{0}_{1 \times d} \quad \text{as } n \to +\infty, \qquad \forall \omega \in \Omega'_{x_0}.$$

Using dichotomy, we show that if Ξ is product bounded, i.e., $\exists \beta > 0$ such that

$$\| \prod\nolimits_{\ell=1}^{n} \xi_\ell(\omega) \|_2 \leq \beta \quad \forall n \geq 1 \text{ and } \mathbb{P}\text{-a.e. } \omega \in \Omega;$$

then Ξ is pointwise convergent if and only if it is pointwise exponentially convergent.

Keywords: Markovian jump linear system, pointwise stabilization, random products of matrices
2010 MSC: Primary 61J10, 93C05; Secondary 37N35, 93D20, 93A30.

1. Introduction

1.1. Motivations

Let $S = \{S_1, \ldots, S_K\}$ be a set that consists of K real $d \times d$ matrices, where K and d both are integers with $2 \leq K < \infty$ and $2 \leq d < \infty$. The system S is said to be

---

- "pointwise convergent" if for each $x \in \mathbb{R}^{1 \times d}$, there is an infinite switching sequence $i.(x) \colon \mathbb{N} \to \{1, \ldots, K\}$ such that $\lim_{n \to +\infty} x \prod_{k=1}^{n} S_{i_k(x)} = \mathbf{0}_{1 \times d}$;

- "pointwise exponentially convergent" if for each $x \in \mathbb{R}^{1 \times d}$, there is a switching sequence $i.(x) \colon \mathbb{N} \to \{1, \ldots, K\}$ such that $\limsup_{n \to +\infty} \frac{1}{n} \log \|x \prod_{k=1}^{n} S_{i_k(x)}\|_2 < 0$.

Here and in the sequel, $\mathbb{N} = \{1, 2, \ldots\}$ is the natural number set, $\mathbf{0}_{d_1 \times d_2}$ stands for the origin of $\mathbb{R}^{d_1 \times d_2}$, and $\| \cdot \|_2$ denotes the usual euclidean norm on $\mathbb{R}^{1 \times d}$ defined by $\|x\|_2 = \sqrt{xx^{\mathrm{T}}}$ for any row-vector $x \in \mathbb{R}^{1 \times d}$.

It is clear that the notion of "convergence" here is abused as it is referring to the usual approach to "convergence" in the stability theory that requires convergence taking place for any (or almost any) switching sequences $i. = (i_n)_{n=1}^{+\infty}$. Here convergence takes place only for some index sequence $i.(x)$.

Further, S is said to be

- "consistently convergent" (also called "uniformly convergent" in, for example, [42, 13]), if the switching sequence $i.(x)$ in the pointwise convergence can be taken independent of the initial state x; that is to say, there exists a switching sequence $i. \colon \mathbb{N} \to \{1, \ldots, K\}$ such that $\lim_{n \to +\infty} \prod_{k=1}^{n} S_{i_k} = \mathbf{0}_{d \times d}$.

These concepts arise and have been studied naturally in the theory of multi-rate sampled-data control systems and multi-modal linear control systems and for some control optimization problems in, for example, [41, 6, 42, 19, 13, 43, 44, 30, 31, 15].

It is worth to remark that instead of the euclidean norm $\| \cdot \|_2$ here can be considered any vector norm on $\mathbb{R}^{1 \times d}$.

In this paper, we consider the random version of the above important concepts driven by stationary Markov chains. Let $(\Omega, \mathscr{F}, \mathbb{P})$ be a probability space and let

$$\Xi = (\xi_n)_{n=1}^{+\infty}, \quad \text{where } \xi_n \colon \Omega \to \{1, \ldots, K\},$$

be a discrete-time, time-homogeneous, stationary $(\boldsymbol{p}, \boldsymbol{P})$-Markov chain, which naturally induces a "Markovian jump linear system" based on S as follows:[1]

$$x_n = x_0 S_{\xi_1(\omega)} \cdots S_{\xi_n(\omega)}, \quad x_0 \in \mathbb{R}^{1 \times d}, n \geq 1, \ \omega \in \Omega.$$

Here $\boldsymbol{p} = (p_1, \ldots, p_K) \in \mathbb{R}^{1 \times K}$ is the initial distribution and $\boldsymbol{P} = [p_{ij}] \in \mathbb{R}^{K \times K}$ is the Markov transition probability matrix of Ξ with $\boldsymbol{p}\boldsymbol{P} = \boldsymbol{p}$.

To any sample $\omega \in \Omega$, there corresponds a switching sequence $\Xi(\omega) = (\xi_n(\omega))_{n=1}^{+\infty}$ that is named a "trajectory" of the Markov chain Ξ in the textbooks of stochastic processes. Moreover, for any switching sequence $i. = (i_n)_{n=1}^{+\infty}$, there need not be some sample ω satisfying $\xi_n(\omega) = i_n$ for all $n \geq 1$. However, in the situation of probability peoples are only concerned with the events of positive-probability. So, motivated by the above notations, we introduce the following concepts.

Definition 1.1. The Markovian jump linear system (S, Ξ) is called to be

---

[1]In almost all available literature, peoples directly consider the linear Markovian process

$$\Xi = (\xi_n)_{n \geq 1}, \quad \text{where } \xi_n \colon \Omega \to S \text{ instead of } \xi_n \colon \Omega \to \{1, \ldots, K\},$$

as formulated in Abstract. This looks more concise. However, our treatment presented here enables us to employ the abstract theory of symbolic dynamical systems. See Section 2 below.

(a) "pointwise convergent", if to any initial state $x_0 \in \mathbb{R}^{1 \times d}$, there corresponds a measurable subset $\Omega_{x_0} \subset \Omega$ with $\mathbb{P}(\Omega_{x_0}) > 0$ such that

$$x_0 \prod_{k=1}^{n} S_{\xi_k(\omega)} \to \mathbf{0}_{1 \times d} \text{ as } n \to +\infty, \quad \forall \omega \in \Omega_{x_0};$$

(b) "pointwise exponentially convergent", if to any initial state $x_0 \in \mathbb{R}^{1 \times d}$, there corresponds a measurable subset $\Omega'_{x_0} \subset \Omega$ with $\mathbb{P}(\Omega'_{x_0}) > 0$ such that

$$\limsup_{n \to +\infty} \frac{1}{n} \log \|x_0 \prod_{k=1}^{n} S_{\xi_k(\omega)}\|_2 < 0, \quad \forall \omega \in \Omega'_{x_0};$$

(c) "consistently exponentially convergent", if there exists a measurable subset $\Omega'' \subset \Omega$ with $\mathbb{P}(\Omega'') > 0$ such that

$$\limsup_{n \to +\infty} \frac{1}{n} \log \|\prod_{k=1}^{n} S_{\xi_k(\omega)}\|_2 < 0, \quad \forall \omega \in \Omega''.$$

Here $\|A\|_2$ denotes the matrix norm induced by the euclidean vector norm $\| \cdot \|_2$ on $\mathbb{R}^{1 \times d}$, for any $A \in \mathbb{R}^{d \times d}$.

Here "consistently" only means that the choice of the driving sample ω is independent of any initial state $x_0 \in \mathbb{R}^{1 \times d}$. We should notice that the consistent exponential convergence of (S, Ξ) over ω is essentially weaker than the so-called "uniform exponential stability" over ω: $\exists C > 0$ and $0 < \gamma < 1$ such that

$$\|\prod_{k=1}^{m} S_{\xi_{\ell+k}(\omega)}\|_2 \leq C\gamma^m \quad \forall \ell \geq 0, m \geq 1.$$

In addition, since Ξ does not need to be irreducible (or equivalently, not need to be ergodic; see Section 2), we cannot require $\mathbb{P}(\Omega'') = 1$ here in general.

It is obvious that (c) $\Rightarrow$ (b) $\Rightarrow$ (a), but not vice versa in general.

Remark 1.2. Because considering a deterministic sample ω makes no sense in probability theory and random/stochastic stability theory, it is necessary to require the property of positive probability: $\mathbb{P}(\Omega_{x_0}) > 0$, $\mathbb{P}(\Omega'_{x_0}) > 0$, and $\mathbb{P}(\Omega'') > 0$, in Definition 1.1. That means all $\Omega_{x_0}, \Omega'_{x_0}$ and Ω'' to be non-negligible events.

1.2. Main results

In this paper, we will mainly show in Section 3 a random stabilizability theorem, stated as follows.

Theorem A. *Let the Markovian jump linear system (S, Ξ) be product bounded; i.e., there exists some $\beta \geq 1$ such that for $\mathbb{P}$-a.e. $\omega \in \Omega$,*

$$\|\prod_{\ell=1}^{n} S_{\xi_\ell(\omega)}\|_2 \leq \beta \quad \forall n \geq 1.$$

Then, (S, Ξ) is pointwise convergent if and only if it is pointwise exponentially convergent.

We notice that if S is irreducible[2] with the joint spectral radius $\hat{\rho}(S) = 1$ (it will be precisely defined in Section 3.1.2), then (S, Ξ) is product bounded; see, e.g., N. Barabanov [3]. The product boundedness condition, also named as "Lyapunov stability" in ODE, is both practically important and academically challenging [32, 1, 23, 2, 45]. Indeed, it is desirable in many practical issues and is closely related to periodic solutions and limit cycles [4, 5].

Such an equivalence theorem will play a key role in creating upper bounds, finding convergence rates and exploiting other basic system properties for Markovian jump linear systems, as done in the deterministic case, for example, in [44, 26].

To prove Theorem A, we need two important tools. One is the ergodic theory of Markov chains established in Section 2. And the other is the following dichotomy decomposition theorem (Theorem B), which will be proved in a more general framework in Section 4 (Theorem B$'$), using the classical multiplicative ergodic theorem [37] and the interesting idea of limit-semigroup due to D.J. Hartfiel [21, 22] and F. Wirth [49].

A vector norm $\|\cdot\|_*$ on $\mathbb{R}^{1 \times d}$ is called a "pre-extremal" norm of S, if its induced matrix norm on $\mathbb{R}^{d \times d}$ is such that $\|S_i\|_* \leq 1$ for all $i \in \{1, \ldots, K\}$. If S is product bounded, then such a pre-extremal norm always exists; see, for example, [11, 28, 18, 36, 23, 15].

Theorem B. *Let S be product bounded, i.e., there exists a universal constant $\beta \geq 1$ such that for any $i \colon \mathbb{N} \to \{1, \ldots, K\}$,*

$$\|\prod_{\ell=1}^{n} S_{\xi_\ell(\omega)}\|_2 \leq \beta \quad \forall n \geq 1.$$

Then to $\mathbb{P}$-a.e. $\omega \in \Omega$, there corresponds a splitting of $\mathbb{R}^{1 \times d}$ into subspaces

$$\mathbb{R}^{1 \times d} = E^s(\omega) \oplus E^c(\omega)$$

such that $\omega \mapsto E^s(\omega)$ is measurable and that

$$\lim_{n \to +\infty} \frac{1}{n} \log \|x_0 S_{\xi_1(\omega)} \cdots S_{\xi_n(\omega)}\|_2 < 0 \qquad \forall x_0 \in E^s(\omega),$$

$$\liminf_{n \to +\infty} \|x_0 S_{\xi_1(\omega)} \cdots S_{\xi_n(\omega)}\|_2 > 0 \qquad \forall x_0 \in \mathbb{R}^{1 \times d} \setminus E^s(\omega),$$

and

$$\|x_0 S_{\xi_1(\omega)} \cdots S_{\xi_n}(\omega)\|_* = \|x_0\|_* \ for \ n \geq 1 \qquad \forall x_0 \in E^c(\omega)$$

for any pre-extremal norm $\|\cdot\|_*$ of S on $\mathbb{R}^{1 \times d}$.

This theorem is of independent interest. It is important, not only to the proof of Theorem A, but also to approximation of the joint spectral radius, almost sure partial stability, and extremal property of orbits, of S; for example, see related works in [3, 36, 29, 33, 15].

On the other hand, if (S, Ξ) is (non-uniformly) hyperbolic over the sample point ω, i.e., the central manifold $E^c(\omega)$ is replaced with the unstable manifold defined as

$$E^u(\omega) = \left\{ x \in \mathbb{R}^{1 \times d} \colon \lim_{n \to +\infty} \frac{1}{n} \log \|x S_{\xi_1(\omega)} \cdots S_{\xi_n(\omega)}\|_2 > 0 \right\} \cup \{\mathbf{0}_{1 \times d}\}$$

---

[2] The matrix family S is said to be "irreducible" if there are no common, proper, nonempty, and invariant subspaces of $\mathbb{R}^{1 \times d}$, for each member of S. See, e.g., [3].

and

$$\mathbb{R}^{1\times d} = E^s(\omega) \oplus E^u(\omega),$$

then it holds trivially that

$$\liminf_{n\to+\infty} \|x_0 \prod_{\ell=1}^{n} \mathbf{S}_{\xi_\ell(\omega)}\|_2 > 0 \quad \forall x_0 \in \mathbb{R}^{1\times d} \setminus E^s(\omega),$$

since in fact

$$\lim_{n\to+\infty} \frac{1}{n} \log \|x_0 \prod_{\ell=1}^{n} \mathbf{S}_{\xi_\ell(\omega)}\|_2 > 0 \quad \forall x_0 \in \mathbb{R}^{1\times d} \setminus E^s(\omega)$$

from the Lyapunov exponent theory.[3]

However, in our situation, some essential difficulties, see for example, Example 4.6 below, come out of the existence of the central manifold $E^c(\omega)$. So, Theorem B might become an intuitive example of systems beyond hyperbolic.

We will end this paper with concluding remarks in Section 5.

2. Preliminary ergodic theory of Markov chains

In this section, we will introduce the grounds of the ergodic theory of stationary Markov chains, which will be used in the proofs of our main results — Theorems A and B, presented in Sections 3 and 4.

2.1. Basic notions

Let $\Xi = (\xi_n)_{n=1}^{+\infty}$ where $\xi_n \colon \Omega \to \mathcal{K}$, be a discrete-time, *time-homogeneous*, *stationary* $(\boldsymbol{p}, \boldsymbol{P})$-Markov chain, defined on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with the finite state-space

$$\mathcal{K} = \{1, \ldots, K\}$$

that is equipped with the discrete topology. Notice here that "time-homogeneity" means that the transition probabilities,

$$\mathbb{P}[\xi_{n+1} = j \,|\, \xi_n = i] = p_{ij} \; \forall i, j \in \mathcal{K} \quad \text{where } \boldsymbol{P} = [p_{ij}] \in \mathbb{R}^{K\times K},$$

all do not depend upon the time n; and "stationary" means $\boldsymbol{p}\boldsymbol{P} = \boldsymbol{p}$. This implies that $\xi_1, \xi_2, \ldots$ are identically distributed random variables. However, they are not necessarily independent. It is easy to see

$$\mathbb{P}[\xi_1 = i_1, \ldots, \xi_n = i_n] = p_{i_1} p_{i_1 i_2} \cdots p_{i_{n-1} i_n}$$

---

[3]Here we only need this simple fact: If two real sequences $\varphi = (\varphi_n)$ and $\psi = (\psi_n)$ are such that

$$\chi(\varphi) := \lim_{n\to+\infty} \frac{1}{n} \log |\varphi_n| \neq \chi(\psi) := \lim_{n\to+\infty} \frac{1}{n} \log |\psi_n|,$$

then $\lim_{n\to+\infty} \frac{1}{n} \log |\varphi_n + \psi_n| = \max\{\chi(\varphi), \chi(\psi)\}$.

for any words

$$(i_1, \ldots, i_n) \in \mathcal{K}^n := \overbrace{\mathcal{K} \times \cdots \times \mathcal{K}}^{n\text{-time}}$$

of finite-length $n \geq 2$. Here the probability vector $\boldsymbol{p} = (p_1, \ldots, p_K) \in \mathbb{R}^{1 \times K}$ is the initial distribution of the Markov chain Ξ, i.e., $\mathbb{P}[\xi_1 = k] = p_k$ for all $1 \leq k \leq K$.

2.2. Induced symbolic dynamical systems

We denote by $\Sigma_{\mathcal{K}}^+$ the set of all infinite switching sequences $i. \colon \mathbb{N} \to \mathcal{K}$. Here it is convenient to place the variables $1, 2, \ldots$ at the subscript position. Then, under the infinite product topology that can be generated by the cylinders

$$[i'_1, \ldots, i'_\ell] = \left\{ i. \in \Sigma_{\mathcal{K}}^+ \mid i_1 = i'_1, \ldots, i_\ell = i'_\ell \right\}$$

for all $\ell \geq 1$ and any words $(i'_1, \ldots, i'_\ell) \in \mathcal{K}^\ell$ of finite-length ℓ, $\Sigma_{\mathcal{K}}^+$ is a compact space as well as the one-sided Markov shift

$$\theta_+ \colon \Sigma_{\mathcal{K}}^+ \to \Sigma_{\mathcal{K}}^+; \quad i. = (i_n)_{n=1}^{+\infty} \mapsto i_{.+1} = (i_{n+1})_{n=1}^{+\infty}$$

is a continuous, surjective transformation.

Then, by the joint random variable

$$\Xi \colon \Omega \to \Sigma_{\mathcal{K}}^+; \quad \omega \mapsto \Xi(\omega) = (\xi_n(\omega))_{n=1}^{+\infty},$$

we can obtain a natural probability distribution, called the "$(\boldsymbol{p}, \boldsymbol{P})$-Markovian measure" and write as $\mu_{\boldsymbol{p}, \boldsymbol{P}}$, on $\Sigma_{\mathcal{K}}^+$, such that

$$\mu_{\boldsymbol{p}, \boldsymbol{P}}([i_1, \ldots, i_n]) = \begin{cases} p_{i_1} & \text{if } n = 1; \\ p_{i_1} p_{i_1 i_2} \cdots p_{i_{n-1} i_n} & \text{if } n \geq 2, \end{cases}$$

for all cylinder sets $[i_1, \ldots, i_n] \subset \Sigma_{\mathcal{K}}^+$.

It should be noted here that $\mu_{\boldsymbol{p}, \boldsymbol{P}}$ is not necessarily equal to the infinite product of the initial distribution of Ξ, for $\xi_1, \xi_2, \ldots$ need not be independent each others.

The following is basic for our arguments later.

Lemma 2.1. *The one-sided Markov shift* $\theta_+ \colon \Sigma_{\mathcal{K}}^+ \to \Sigma_{\mathcal{K}}^+$ *preserves the* $(\boldsymbol{p}, \boldsymbol{P})$-*Markovian measure* $\mu_{\boldsymbol{p}, \boldsymbol{P}}$; *that is to say,*

$$\mu_{\boldsymbol{p}, \boldsymbol{P}}(B) = \mu_{\boldsymbol{p}, \boldsymbol{P}} \circ \theta_+^{-1}(B)$$

for all Borel subsets $B \subset \Sigma_{\mathcal{K}}^+$.

Proof. Since the Borel σ-field $\mathscr{F}_{\Sigma_{\mathcal{K}}^+}$ of $\Sigma_{\mathcal{K}}^+$ may be generated by all the cylinder sets $[i_1, \ldots, i_n]$, we only need to prove

$$\mu_{\boldsymbol{p}, \boldsymbol{P}}([i_1, \ldots, i_n]) = \mu_{\boldsymbol{p}, \boldsymbol{P}}(\theta_+^{-1}[i_1, \ldots, i_n]).$$

In fact, by definition, we have

$$
\begin{aligned}
\mu_{\boldsymbol{p},\boldsymbol{P}}([i_1,\ldots,i_n]) &= p_{i_1} p_{i_1 i_2} \cdots p_{i_{n-1} i_n} \\
&= \sum_k p_k p_{k i_1} p_{i_1 i_2} \cdots p_{i_{n-1} i_n} \\
&= \sum_k \mu_{\boldsymbol{p},\boldsymbol{P}}([k, i_1,\ldots,i_n]) \\
&= \mu_{\boldsymbol{p},\boldsymbol{P}} \circ \theta_+^{-1}([i_1,\ldots,i_n])
\end{aligned}
$$

for all words $(i_1,\ldots,i_n) \in \mathcal{K}^n$ and any $n \geq 1$. This completes the proof of Lemma 2.1. $\quad\square$

This simple result induces an affirmative answer to this question: Is $(\xi_{\ell+n}(\omega))_{n\geq 1}$, for any ω and $\ell \geq 1$, a trajectory of the Markov process Ξ? In fact, from the discrete topology of $\mathcal{K}^n$ and

$$
\Xi(\Omega) = \bigcap_{n=1}^{+\infty} \{(\xi_1(\omega),\ldots,\xi_n(\omega)) \colon \omega \in \Omega\} \times \mathcal{K} \times \mathcal{K} \times \cdots ,
$$

we see $\Xi(\Omega)$ is a Borel subset of $\Sigma_{\mathcal{K}}^+$ with $\mu_{\boldsymbol{p},\boldsymbol{P}}(\Xi(\Omega)) = 1$; if we let

$$
\Sigma = \bigcap_{\ell=0}^{+\infty} \theta_+^{-\ell} \Xi(\Omega) \quad \text{and} \quad \Omega_\infty = \Xi^{-1}(\Sigma);
$$

then $\mu_{\boldsymbol{p},\boldsymbol{P}}(\Sigma) = 1$ and so $\mathbb{P}(\Omega_\infty) = 1$; it is easy to check that for any $\omega \in \Omega_\infty$ and any $\ell \geq 1$, $(\xi_{\ell+n}(\omega))_{n\geq 1}$ is still a trajectory of Ξ, i.e., there is some other sample point $\omega' \in \Omega_\infty$ such that $(\xi_n(\omega'))_{n\geq 1} = (\xi_{\ell+n}(\omega))_{n\geq 1}$.

Recall from P. Walters [48] that an invariant probability measure μ of the shift θ_+ on $\Sigma_{\mathcal{K}}^+$ is called "ergodic" if for $B \in \mathscr{F}_{\Sigma_{\mathcal{K}}^+}$, the following equality

$$
\mu\left((\theta_+^{-1}(B) \setminus B) \cup (B \setminus \theta_+^{-1}(B))\right) = 0
$$

implies that $\mu(B) = 1$ or 0.

Then there is a well-known fact.

Lemma 2.2. *The $(\boldsymbol{p},\boldsymbol{P})$-Markovian probability $\mu_{\boldsymbol{p},\boldsymbol{P}}$ is θ_+-ergodic on $\Sigma_{\mathcal{K}}^+$ if and only if the transition matrix $\boldsymbol{P}$ is irreducible.*

Here the transition probability matrix $\boldsymbol{P}$ is called "irreducible" if for any pair $i, j \in \mathcal{K}$, there is some $n = n(i, j) \geq 1$ such that $p_{ij}^{(n)} > 0$, where $p_{ij}^{(n)}$ is the (i, j)-th element of the n-time product matrix $\boldsymbol{P}^\ell$. It is worth to mention here that this "irreducibility" has nothing in common with "irreducibility" explained to a family of matrices in Footnote 2.

2.3. Ergodic decomposition of Markovian probability

Since the Markov transition probability matrix $\boldsymbol{P}$ is not necessarily irreducible in our situation, we need to consider the ergodic decomposition of the $(\boldsymbol{p},\boldsymbol{P})$-Markovian probability $\mu_{\boldsymbol{p},\boldsymbol{P}}$.

Hereafter, assume $\boldsymbol{p} > 0$, i.e., $p_k > 0 \;\forall k \in \mathcal{K}$; otherwise, we only need to replace the state-space $\mathcal{K}$ of the Markov chain Ξ with $\mathcal{K} \setminus \{k\}$ if $p_k = 0$ for some $1 \leq k \leq K$.

A state $k \in \mathcal{K}$ is called "recurrent" for Ξ, if the conditional probability

$$\mathbb{P}[\omega \in \Omega : \exists n_\ell \nearrow +\infty \text{ such that } \xi_{n_\ell}(\omega) = k \,|\, \xi_1 = k] = 1.$$

If $k \in \mathcal{K}$ is not recurrent, then it is called "non-recurrent". Two states k_1, k_2, each accessible to the other, i.e, $p_{k_1 k_2}^{(m)} > 0$ and $p_{k_2 k_1}^{(n)} > 0$ for some pair $m, n \geq 1$, are said to "communicate" and we write $k_1 \leftrightsquigarrow k_2$. The concept of $\leftrightsquigarrow$ is an equivalence relation.

Then according to the classical theory of stochastic processes, for example, [12], there exists the following basic partition of the totality of states:

$$\mathcal{K} = \mathcal{K}_0 \cup \mathcal{K}_1 \cup \cdots \cup \mathcal{K}_r$$

such that

- $\mathcal{K}_0$ consists of all the non-recurrent states of the Markov chain Ξ;

- each $\mathcal{K}_i, 1 \leq i \leq r$, is closed and communicative, i.e., for any $k, k' \in \mathcal{K}_i$ and $k'' \notin \mathcal{K}_i$, we have

$$k \leftrightsquigarrow k', \quad p_{kk''}^{(n)} = 0 \ \forall n \geq 1.$$

Then, based on each component $\mathcal{K}_i, 1 \leq i \leq r$, one can define a symbolic system $\theta_+ : \Sigma_{\mathcal{K}_i}^+ \to \Sigma_{\mathcal{K}_i}^+$, where $\Sigma_{\mathcal{K}_i}^+ = \{i. : \mathbb{N} \to \mathcal{K}_i\}$. It is easily seen that $\Sigma_{\mathcal{K}_i}^+$ is closed invariant subspace of $\Sigma_{\mathcal{K}}^+$.

On the other hand, there hold the following two basic results.

Lemma 2.3. $\mu_{p,P}(\Sigma_{\mathcal{K}_i}^+) > 0$, for each $1 \leq i \leq r$.

Proof. By the closedness of the component $\mathcal{K}_i$ in the basic partition, it is easy to see

$$\mu_{p,P}(\Sigma_{\mathcal{K}_i}^+) = \sum_{k \in \mathcal{K}_i} p_k > 0$$

from the definition of $\mu_{p,P}$. This completes the proof of Lemma 2.3. ∎

Lemma 2.4. $\mathcal{K}_0 = \varnothing$ under the assumption $p > 0$. In general case, $\mathcal{K}_0 = \{k : p_k = 0\}$.

Proof. Suppose, by contradiction, that $\mathcal{K}_0 \neq \varnothing$. Let $k \in \mathcal{K}_0$. Then, for the cylinder $[k]$ we have $\mu_{p,P}([k]) = p_k > 0$. Applying Poincaré's Recurrence Theorem ([48, Theorem 1.4], also see Section 4.1 below) to $(\Sigma_{\mathcal{K}}^+, \theta_+, \mu_{p,P})$, we see that there exists a Borel set $F \subset [k]$ with $\mu_{p,P}(F) = p_k$ such that for each $i. \in F$ there is a sequence $n_1 < n_2 < \cdots$ of natural numbers satisfying $i._{+n_\ell} \in F$, i.e, $i_{n_\ell + 1} = k$, for each $\ell \geq 1$. This implies that k would be a recurrent state for Ξ, a contradiction. This proves Lemma 2.4. ∎

Let $\alpha_i = \mu_{p,P}(\Sigma_{\mathcal{K}_i}^+)$ for $1 \leq i \leq r$. Then $0 < \alpha_i \leq 1$ and $\alpha_1 + \cdots + \alpha_r = 1$. So,

$$\Sigma_{\mathcal{K}}^+ = \Sigma_{\mathcal{K}_1}^+ \cup \cdots \cup \Sigma_{\mathcal{K}_r}^+ \quad (\mathrm{mod} \, \mu_{p,P})$$

is a measurable, not topological, partition of $\Sigma_{\mathcal{K}}^+$. Define conditional θ_+-invariant probability measures $\mu_{p,P}(\cdot | \mathcal{K}_i)$ on $\Sigma_{\mathcal{K}_i}^+$ by

$$\mu_{p,P}(B | \mathcal{K}_i) = \frac{\mu_{p,P}(B \cap \Sigma_{\mathcal{K}_i}^+)}{\alpha_i} \quad \forall B \in \mathscr{F}_{\Sigma_{\mathcal{K}}^+},$$

for each $1 \le i \le r$. Then,

$$\mu_{p,P}(\cdot) = \alpha_1 \mu_{p,P}(\cdot \,|\, \mathcal{K}_1) + \cdots + \alpha_r \mu_{p,P}(\cdot \,|\, \mathcal{K}_r).$$

Next, we will show that this is just the ergodic decomposition of $\mu_{p,P}$.

Theorem 2.5. *For each $1 \le i \le r$, $\mu_{p,P}(\cdot \,|\, \mathcal{K}_i)$ is an ergodic probability measure of θ_+, supported on the subspace $\Sigma_{\mathcal{K}_i}^+$.*

Proof. Let $\Omega_i = \{\omega \in \Omega \colon \Xi(\omega) \in \Sigma_{\mathcal{K}_i}^+\}$ for each $1 \le i \le r$. Then,

$$\Omega = \Omega_1 \cup \cdots \cup \Omega_r \quad (\mathrm{mod}\ \mathbb{P})$$

is a measurable partition of Ω. Moreover, $\Xi_{|\Omega_i} = (\xi_{n|\Omega_i})_{n=1}^{+\infty}$ is a time-homogeneous, stationary Markov chain, defined on the conditional probability space $(\Omega_i, \mathscr{F}_{|\Omega_i}, \mathbb{P}(\cdot \,|\, \Omega_i))$ with the state space $\mathcal{K}_i$. Clearly, its initial distribution is

$$p_i = (p_{k_1}, \ldots, p_{k_\ell})/\alpha_i \quad \text{if } \mathcal{K}_i = \{k_1, \ldots, k_\ell\}$$

and its transition probability matrix is

$$P_i = [p_{ij}] \in \mathbb{R}^{\mathcal{K}_i \times \mathcal{K}_i}$$

which is irreducible from the closedness of the component $\mathcal{K}_i$ in the basic partition of $\mathcal{K}$. Thus from Lemma 2.2, it follows that $\mu_{p,P}(\cdot \,|\, \mathcal{K}_i) = \mu_{p_i,P_i}$ is ergodic of θ_+.

This completes the proof of Theorem 2.5. □

This ergodic decomposition will provide us convenience for proving our main results below.

Remark 2.6. Since all the θ_+-invariant probabilities form a compact convex set $\mathcal{M}_{\mathrm{inv}}(\Sigma_{\mathcal{K}}^+, \theta_+)$ and all the θ_+-ergodic probabilities are just its extreme points, we can directly use the Choquet representation theorem to express each member of $\mathcal{M}_{\mathrm{inv}}(\Sigma_{\mathcal{K}}^+, \theta_+)$ in terms of ergodic members. See R. Phelps [38]. Hence every $\mu \in \mathcal{M}_{\mathrm{inv}}(\Sigma_{\mathcal{K}}^+, \theta_+)$ is a generalised convex combination of ergodic probabilities. However, it is important that a (p, P)-Markovian probability $\mu_{p,P}$ has only finitely many ergodic components and each component is still a Markovian probability from Theorem 2.5.

3. Pointwise and consistent stabilizability

This section will be mainly devoted to proving Theorem A stated in Section 1. In addition, we will prove some equivalence between consistent stabilizations.

Let $S = \{S_1, \ldots, S_K\} \subset \mathbb{R}^{d \times d}$ be arbitrarily given K matrices and $\mathcal{K} = \{1, \ldots, K\}$, where K and d both ≥ 2. Recall that S is called "product bounded" if the multiplicative semigroup S^+, generated by S, is bounded in $\mathbb{R}^{d \times d}$. This is equivalent to that there exists a constant $\beta > 0$ such that for any $i. \in \Sigma_{\mathcal{K}}^+$,

$$\|S_{i_1} \cdots S_{i_n}\|_2 \le \beta \quad \forall n \ge 1.$$

We will study the random stabilizability of S, driven by a discrete-time Markov chain.

Let $\Xi = (\xi_n)_{n=1}^{+\infty}$, where $\xi_n \colon (\Omega, \mathscr{F}, \mathbb{P}) \to \mathcal{K}$, be a discrete-time, time-homogeneous, stationary (p, P)-Markov chain, as described in Section 1.

3.1. Pointwise stabilizability: an abstraction of Theorem A

To prove Theorem A, we consider an abstraction version. For any θ_+-invariant probability measure μ on $\Sigma_{\mathcal{K}}^+$, we say (S, μ) to be

(a)′ "pointwise convergent", if to any initial state $x_0 \in \mathbb{R}^{1 \times d}$, there corresponds a Borel subset $\Sigma_{x_0} \subset \Sigma_{\mathcal{K}}^+$ with $\mu(\Sigma_{x_0}) > 0$ such that

$$x_0 \prod_{\ell=1}^{n} S_{i_\ell} \to \mathbf{0}_{1 \times d} \quad \text{as} \quad n \to +\infty, \quad \forall i. \in \Sigma_{x_0};$$

(b)′ "pointwise exponentially convergent", if to any initial state $x_0 \in \mathbb{R}^{1 \times d}$, there corresponds a Borel subset $\Sigma'_{x_0} \subset \Sigma_{\mathcal{K}}^+$ with $\mu(\Sigma'_{x_0}) > 0$ such that

$$\limsup_{n \to +\infty} \frac{1}{n} \log \|x_0 \prod_{\ell=1}^{n} S_{i_\ell}\|_2 < 0, \quad \forall i. \in \Sigma'_{x_0}.$$

Then from Section 2.2 and considering the (p, P)-Markovian probability $\mu_{p,P}$, it follows that the statement of Theorem A is true if and only if there holds the following result:

Theorem 3.1. *Let S be product bounded. Then, $(S, \mu_{p,P})$ is pointwise convergent if and only if it is pointwise exponentially convergent.*

It should be interesting to notice that for Theorem 3.1, we can require that $\Sigma'_{x_0} \subset \Sigma_{x_0}$ with $\mu_{p,P}(\Sigma'_{x_0}) = \mu_{p,P}(\Sigma_{x_0})$. And a similar requirement can be satisfied for Theorem A.

We also should note that the deterministic version of Theorem 3.1 was proven by Z. Sun [44, Theorem 1]. However, our random version obtained here cannot be derived from Sun's theorem and approach because of the requirement of positive probability in Definition 1.1.

Proof of Theorem A. We only need to prove the necessity. Let (S, Ξ) be pointwise convergent. Then from Theorem 2.5, it follows that there exists at least one ergodic component $\mu_{p,P}(\cdot|\mathcal{K}_i)$ of $\mu_{p,P}$ such that $(S, \mu_{p,P}(\cdot|\mathcal{K}_i))$ is pointwise convergent. Since (S, Ξ) is product bounded, one can find a constant $\beta > 0$ such that

$$\|S_{i_1} \cdots S_{i_n}\|_2 \le \beta \quad \forall n \ge 1$$

for $\mu_{p,P}(\cdot|\mathcal{K}_i)$-a.e. $i. \in \Sigma_{\mathcal{K}}^+$. On the other hand, the density points of $\mu_{p,P}(\cdot|\mathcal{K}_i)$ are dense in the subspace $\Sigma_{\mathcal{K}_i}^+$; so, S is product bounded over $\Sigma_{\mathcal{K}_i}^+$. Now, Theorem A follows from Theorem 3.1 with replacing $\mu_{p,P}$ by $\mu_{p,P}(\cdot|\mathcal{K}_i)$. $\square$

We need to note here that the product boundedness of (S, Ξ) is weaker than that of S over $\Sigma_{\mathcal{K}}^+$ in general, unless the transition probability matrix P is irreducible.

3.1.1. Proof of Theorem 3.1

For any θ_+-invariant measure μ, the pointwise exponential convergence of (S, μ) implies obviously the pointwise convergence. Thus, according to the ergodic decomposition (Theorem 2.5), Theorem 3.1 follows immediately from the following statement.

Proposition 3.2. *Let S be product bounded and μ an ergodic probability measure of θ_+ on $\Sigma_{\mathcal{K}}^+$. If (S, μ) is pointwise convergent, then it is pointwise exponentially convergent.*

Proof. This statement comes at once from Theorem B that is stated in Section 1 and will be proved in Section 4. In fact, let $x_0 \in \mathbb{R}^{1 \times d} \setminus \{\mathbf{0}_{1 \times d}\}$. Since (S, μ) is pointwise convergent, one can find some Borel subset $\Sigma_{x_0} \subset \Sigma_{\mathcal{K}}^+$ with $\mu(\Sigma_{x_0}) > 0$ such that $\|x_0 S_{i_1} \cdots S_{i_n}\|_2 \to 0$ as $n \to +\infty$ for all $i. \in \Sigma_{x_0}$. According to Theorem B (precisely, since here μ is more general than $\mu_{p,P}$, we need Theorem B′ and Weak Birkhoff Recurrence Theorem stated in Section 4), one can further choose a Borel subset $\Sigma'_{x_0} \subseteq \Sigma_{x_0}$ with $\mu(\Sigma'_{x_0}) > 0$ such that $x_0 \in E^s(i.)$ for any $i. \in \Sigma'_{x_0}$.

This completes the proof of Proposition 3.2. □

Hence, the proof of Theorem A is completed if we recognize the statement of Theorem B.

3.1.2. *A further question related to Theorem A*

To describe the maximal growth rate of the trajectories generated by random products of matrices $S_1, \ldots, S_K$ in S, in [40] G.-C. Rota and G. Strang introduced the very important concept – *joint spectral radius of S* – by

$$\hat{\rho}(S) = \limsup_{n \to +\infty} \left\{ \max_{i. \in \Sigma_{\mathcal{K}}^+} \|S_{i_1} \cdots S_{i_n}\|_2^{1/n} \right\}.$$

It is well known that $\hat{\rho}(S) < 1$ if and only if S is absolutely (uniformly) exponentially stable, i.e.,

$$\lim_{n \to +\infty} \prod_{\ell=1}^{n} S_{i_\ell} = \mathbf{0}_{d \times d} \quad \forall i. \in \Sigma_{\mathcal{K}}^+;$$

see N. Barabanov [3]. We notice that if S is product bounded or, more generally, polynomially bounded in $\mathbb{R}^{d \times d}$ as in L. Gurvits and L. Rodman [19], then $\hat{\rho}(S) \leq 1$. As shown by the example

$$S = \left\{ \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \right\},$$

$\hat{\rho}(S) = 1$ need not imply the product boundedness of S.

Let $\Xi = (\xi_n)_{n=1}^{+\infty}$, where $\xi_n \colon (\Omega, \mathscr{F}, \mathbb{P}) \to \mathcal{K}$, be a discrete-time, time-homogeneous, stationary (p, P)-Markov chain as before. Here we ask the following.

Question 3.3. If S is reducible with $\hat{\rho}(S) = 1$ and (S, Ξ) is pointwise convergent then, is (S, Ξ) pointwise exponentially convergent?

3.2. *Periodically switched stable systems*

For $S = \{S_1, \ldots, S_K\} \subset \mathbb{R}^{d \times d}$, it is called "periodically switched stable" [39, 47, 16] if for any finite-length words $(k_1, \ldots, k_n) \in \mathcal{K}^n$ and $n \geq 1$, the spectral radius $\rho(S_{k_1} \cdots S_{k_n}) < 1$, i.e., over any periodical switching sequences

$$i. = (k_1, \ldots, k_n, k_1, \ldots, k_n, k_1, \ldots, k_n, \ldots), \quad i.e., \ i_{j+\ell n} = k_j \ \forall 1 \leq j \leq n \text{ and } \ell \geq 0,$$

we have $\| \prod_{\ell=1}^{n} S_{i_\ell} \|_2 \to 0$ as $n \to +\infty$.

There are counterexamples which show that the periodical-switched stability need not imply the absolute asymptotic stability of S, namely, $\| \prod_{\ell=1}^{n} S_{i_\ell} \|_2 \to 0$ as $n \to +\infty$ for all $i. \in \Sigma_{\mathcal{K}}^+$. See [10], also [9, 29, 20]. However, in [16, Main Theorem], the authors proved that S is exponentially stable $\mu_{p,P}$-almost surely, if the transition probability matrix P is irreducible, i.e., $\mu_{p,P}$ is ergodic for θ_+.

From the ergodic decomposition theorem (Theorem 2.5) and [16, Main Theorem], we can easily obtain a more general result as follows:

Theorem 3.4. *Let S be periodically switched stable. Then the Markovian jump linear system (S, Ξ) is exponentially stable $\mathbb{P}$-almost surely; that is to say,*

$$\|S_{\xi_1(\omega)} \cdots S_{\xi_n(\omega)}\|_2 \xrightarrow{\text{exponentially fast}} 0 \quad \text{as } n \to +\infty,$$

for $\mathbb{P}$-a.e. $\omega \in \Omega$.

This theorem generalizes the statement (1) of [16, Main Theorem] from ergodic probability case to invariant probability case.

3.3. Consistent stabilizability

Recall, for instance from [42, 43] in the deterministic situation, that $S = \{S_1, \ldots, S_K\} \subset \mathbb{R}^{d \times d}$ is called to be

- "consistently convergent" if there is a switching sequence $i. \colon \mathbb{N} \to \mathcal{K}$ such that

$$\prod_{\ell=1}^{n} S_{i_\ell} \to \mathbf{0}_{d \times d}, \text{ or equivalently, } x \prod_{\ell=1}^{n} S_{i_\ell} \to \mathbf{0}_{1 \times d} \ \forall x \in \mathbb{R}^{1 \times d} \quad \text{as } n \to +\infty;$$

- "consistently exponentially convergent" if there is a switching sequence $i. \colon \mathbb{N} \to \mathcal{K}$ such that

$$\limsup_{n \to +\infty} \frac{1}{n} \log \|\prod_{\ell=1}^{n} S_{i_\ell}\|_2 < 0, \quad \text{i.e.,} \quad \prod_{\ell=1}^{n} S_{i_\ell} \xrightarrow{\text{exponentially fast}} \mathbf{0}_{d \times d} \text{ as } n \to +\infty.$$

The random versions of these concepts driven by the Markov chain $\Xi = (\xi_n)_{n \geq 1}$ can be formulated as follows. The pair (S, Ξ) is called to be

- "consistently convergent" if there is measurable set $\Omega' \subset \Omega$ with $\mathbb{P}(\Omega') > 0$ such that

$$\prod_{\ell=1}^{n} S_{\xi_\ell(\omega)} \to \mathbf{0}_{d \times d} \text{ as } n \to +\infty, \quad \forall \omega \in \Omega';$$

- "consistently exponentially convergent" if there exists a measurable set $\Omega'' \subset \Omega$ with $\mathbb{P}(\Omega'') > 0$ such that

$$\prod_{\ell=1}^{n} S_{\xi_\ell(\omega)} \xrightarrow{\text{exponentially fast}} \mathbf{0}_{d \times d} \text{ as } n \to +\infty, \quad \forall \omega \in \Omega''.$$

In D.P. Stanford and J.M. Urbano [42, Theorem 3.5], it is proved that S is consistently convergent if and only if it is consistently exponentially convergent; more precisely, S is consistently convergent if and only if there is finite-length word $w = (k_1, \ldots, k_m) \in \mathcal{K}^m$ for some $m \geq 1$ such that the spectral radius $\rho(S_{k_1} \cdots S_{k_m}) < 1$. Also see Z. Sun [43, Proposition 4] and J.-W. Lee and G.E. Dullerud [31, Theorem 2].

We notice that although the consistent exponential convergence of S implies, from Y. Huang et al. [25], that there exists some other (p', P')-Markovian probability $\mu_{p', P'}$ such that $(S, \mu_{p', P'})$ is consistently exponentially convergent, yet it cannot imply the consistent exponential convergence of (S, Ξ) in general. This is because $\mu_{p', P'}$, constructed in [25] there, need not equal $\mu_{p, P}$ that has been presented in our situation, and the set of all periodical switching sequences in $\Sigma_{\mathcal{K}}^+$ has $\mu_{p, P}$-measure 0 in general case; see for example, [16]. However, based on the recent work of X. Dai [15] we can obtain the following equivalence result.

Theorem 3.5. *Let $\hat{\rho}(S) = 1$. Then, (S, Ξ) is consistently convergent if and only if it is consistently exponentially convergent.*

Proof. From the ergodic theory presented in Section 2, we only need to prove this claim: $(S, \mu_{p,P})$ *is consistently convergent if and only if it is consistently exponentially convergent.*

Assume $(S, \mu_{p,P})$ is consistently convergent. Then by Theorem 2.5, there exists at least one ergodic component, say $\mu_{p,P}(\cdot | \mathcal{K}_i)$, of $\mu_{p,P}$ such that one can find a Borel set $\Sigma' \subset \Sigma_{\mathcal{K}_i}^+$ with $\mu_{p,P}(\Sigma' | \mathcal{K}_i) > 0$ satisfying $S_{i_1} \cdots S_{i_n} \to \mathbf{0}_{d \times d}$ as $n \to +\infty$ for all $i. \in \Sigma'$. Then from [42, Theorem 3.5], it follows that there is a finite-length word $w = (k_1, \ldots, k_m) \in \mathcal{K}_i^m$ such that $\rho(S_{k_1} \cdots S_{k_m}) < 1$. On the other hand, since the component $\mathcal{K}_i$ is closed and communicative, each point of $\Sigma_{\mathcal{K}_i}^+$ is a density point of $\mu_{p,P}(\cdot | \mathcal{K}_i)$, and hence the support of $\mu_{p,P}(\cdot | \mathcal{K}_i)$, that consists of all its density points, is equal to $\Sigma_{\mathcal{K}_i}^+$.

Therefore, it follows from [15, Theorem C$'$] that for $\mu_{p,P}(\cdot | \mathcal{K}_i)$-a.e. $i. \in \Sigma_{\mathcal{K}_i}^+$,

$$\prod_{\ell=1}^{n} S_{i_\ell} \xrightarrow{\text{exponentially fast}} \mathbf{0}_{d \times d} \quad \text{as } n \to +\infty.$$

This implies that there exists a Borel set $\Sigma'' \subset \Sigma_{\mathcal{K}}^+$ with $\mu_{p,P}(\Sigma'') > 0$ such that

$$\prod_{\ell=1}^{n} S_{i_\ell} \xrightarrow{\text{exponentially fast}} \mathbf{0}_{d \times d} \text{ as } n \to +\infty \quad \forall i. \in \Sigma''.$$

So, $(S, \mu_{p,P})$ is consistently exponentially convergent.

This completes the proof of Theorem 3.5. $\qquad\qquad\square$

Theorem 3.5 is the random version of [42, Theorem 3.5]. We conjecture that this statement still holds without the condition $\hat{\rho}(S) = 1$; that is the following.

Conjecture 3.6. (S, Ξ) is consistently convergent if and only if it is consistently exponentially convergent.

It can be shown by the example $S = \left\{ \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & 2 \end{bmatrix}, \begin{bmatrix} \sqrt{3}/2 & 1/2 \\ -1/2 & \sqrt{3}/2 \end{bmatrix} \right\}$ that the pointwise and consistent convergence of S are not necessarily the same. However, both types of convergence of S are equivalent, for the case that S consists of diagonal matrices (proven in D.P. Stanford and J.M. Urbano [42]) and for the case that S^+ is polynomially bounded (proven in L. Gurvits and L. Rodman [19]).

As a result of Theorem 3.5, we can obtain the following simple random version.

Corollary 3.7. *Let S consist of diagonal matrices with $\hat{\rho}(S) = 1$. Then, (S, Ξ) is pointwise convergent if and only if it is consistently exponentially convergent.*

Proof. Assume (S, Ξ) is pointwise convergent. Then for $x_0 = (1, \ldots, 1) \in \mathbb{R}^{1 \times d}$, there exists a measurable set $\Omega_{x_0} \subset \Omega$ with $\mathbb{P}(\Omega_{x_0}) > 0$ such that

$$x_0 \prod_{\ell=1}^{n} S_{\xi_\ell(\omega)} \to \mathbf{0}_{1 \times d} \text{ as } n \to +\infty \quad \forall \omega \in \Omega_{x_0}.$$

Since S is diagonal, there follows immediately that

$$\prod_{\ell=1}^{n} S_{\xi_\ell(\omega)} \to \mathbf{0}_{d \times d} \text{ as } n \to +\infty \quad \forall \omega \in \Omega_{x_0}.$$

So, (S, Ξ) is consistently convergent and further consistently exponentially convergent from Theorem 3.5.

This proves the statement of Corollary 3.7. ▢

4. Partial stability over almost sure switching sequences

In this section, we will prove Theorem B in a more general framework that is of independent interest for stability analysis of linear switched systems. Hereafter, let

$$S \colon \mathcal{I} \to \mathbb{R}^{d \times d}; \quad i \mapsto S_i$$

be a continuous matrix-valued function defined on a separable metrizable space $\mathcal{I}$, where we assume $2 \le d < +\infty$ and $\mathrm{Card}(\mathcal{I}) \ge 2$.

We consider the stability and stabilization of the discrete-time linear inclusion/control dynamics naturally induced by S:

$$x_n \in \{x_{n-1}S_i\}_{i \in \mathcal{I}}, \quad x_0 \in \mathbb{R}^{1 \times d} \text{ and } n \ge 1.$$

As in [16, 15], we denote by

$$\Sigma_{\mathcal{I}}^+ = \{i_\cdot \colon \mathbb{N} \to \mathcal{I}\}$$

the set of all admissible control/switching sequences of S, equipped with the standard infinite product topology. For any input $(x_0, i_\cdot)$, where $x_0 \in \mathbb{R}^{1 \times d}$ is an initial state and $i_\cdot = (i_n)_{n=1}^{+\infty} \in \Sigma_{\mathcal{I}}^+$ a control/switching sequence, there is a unique output $(x_n(x_0, i_\cdot))_{n=1}^{+\infty}$, called an orbit of S, which corresponds to the unique solution of the linear switched dynamics, written also as S,

$$x_n = x_0 S_{i_1} \cdots S_{i_n}, \quad x_0 \in \mathbb{R}^{1 \times d} \text{ and } n \ge 1$$

over the switching sequence $i_\cdot = (i_n)_{n=1}^{+\infty}$. If the joint spectral radius of S, introduced by G.-C. Rota and G. Strang in [40],

$$\hat{\rho}(S) = \limsup_{n \to +\infty} \sup_{i_\cdot \in \Sigma_{\mathcal{I}}^+} \sqrt[n]{\|S_{i_1} \cdots S_{i_n}\|_2},$$

is less than 1, then $\|x_n(x_0, i_\cdot)\|_2 \to 0$ as $n \to +\infty$ for all pairs $(x_0, i_\cdot) \in \mathbb{R}^{1 \times d} \times \Sigma_{\mathcal{I}}^+$ directly from the definition. If $\hat{\rho}(S) > 1$ then from [17], for almost every (related to some extremal probability) input $(x_0, i_\cdot)$, $\|x_n(x_0, i_\cdot)\|_2$ diverges exponentially fast to $+\infty$ as $n \to +\infty$. Therefore, the most interesting and complicated case is the "neutral type" that holds $\hat{\rho}(S) = 1$.

For the neutral type, N. Barabanov [3] proved that if S is irreducible and finite, then one can find an "extremal norm"[4] $\|\cdot\|_*$ on $\mathbb{R}^{1 \times d}$ such that for any $x_0 \in \mathbb{R}^{1 \times d}$, there corresponds a switching sequence $i_\cdot(x_0) = (i_n(x_0))_{n=1}^{+\infty}$ with

$$\|x_0 S_{i_1(x_0)} \cdots S_{i_n(x_0)}\|_* = \|x_0\|_* \quad \forall n \ge 1.$$

Such an orbit $(x_n(x_0, i_\cdot(x_0)))_{n=1}^{+\infty}$ is called a "$\|\cdot\|_*$-extremal orbit" of S. In V. Kozyakin [29], it is proved that if S is irreducible and finite, then there is an orbit, or, equivalently an input $(x_0, i_\cdot)$,

---

[4] According to [40] and also see [3, 49], for any $S = \{S_i\}_{i \in \mathcal{I}} \subset \mathbb{R}^{d \times d}$, a norm $\|\cdot\|_*$ on $\mathbb{R}^{1 \times d}$ is called "extremal" for S, if $\hat{\rho}(S) = \sup_{i \in \mathcal{I}} \|S_i\|_*$. In the neutral type case, we have $1 = \sup_{i \in \mathcal{I}} \|S_i\|_*$ for any extremal norm $\|\cdot\|_*$ of S.

which is extremal in the sense of all extremal norms $\|\cdot\|_*$ of S; see Corollary 4.1 below. In [10, 9, 29, 33], all these works, for the case of positive 2×2 matrices of a special form, it was shown that "extremal orbits" are generated by the so-called Sturmian sequences. And this was the key point in all the proofs there. In [50, 29, 34], F. Fabian, V. Kozyakin and Ian D. Morris further studied the topological structures of the extremal-norm and the Barabanov-norm sets of S, respectively.

However, in all [40, 3, 10, 9, 50, 29, 33, 34] and elsewhere, there are no descriptions of the structure of extremal control/switching sequences of S from the viewpoint of probability (Corollary 4.1). Conversely, if $i. = (i_n)_{n=1}^{+\infty}$ is $\|\cdot\|_*$-extremal for S in the neutral type case, i.e., $\|S_{i_1}\cdots S_{i_n}\|_* = 1 \ \forall n \geq 1$, what is the topological structure of the corresponding $\|\cdot\|_*$-extremal initial-state set $\{x_0 \in \mathbb{R}^{1\times d} : (x_n(x_0, i.))_{n\geq 1}$ is $\|\cdot\|_*$-extremal for $S\}$? These are important for study of the asymptotic stability and stabilization of the dynamical system S.

For this aim, we need to consider the classical one-sided Markov shift transformation defined as

$$\theta_+ : \Sigma_{\mathcal{I}}^+ \to \Sigma_{\mathcal{I}}^+; \quad i. = (i_n)_{n=1}^{+\infty} \mapsto i._{+1} = (i_{n+1})_{n=1}^{+\infty}.$$

Under the product topology, θ_+ is continuous and surjective, but not one-to-one.

For the case where S is product bounded in $\mathbb{R}^{d\times d}$, it has been shown in [15, Theorem A′] that for any θ_+-ergodic probability μ on $\Sigma_{\mathcal{I}}^+$, either

$$\|S_{i_1}\cdots S_{i_n}\|_2 \xrightarrow{\text{exponentially fast}} 0 \ \text{ as } n \to +\infty \qquad\qquad \text{for } \mu\text{-a.e. } i. \in \Sigma_{\mathcal{I}}^+,$$

or for any norm $\|\cdot\|_*$ of S with $\|S_i\|_* \leq 1$ we have

$$\|S_{i_{k+1}}\cdots S_{i_{k+n}}\|_* = 1 \ \forall k \geq 0 \text{ and } n \geq 1 \qquad\qquad \text{for } \mu\text{-a.e. } i. \in \Sigma_{\mathcal{I}}^+.$$

However, there are no available criteria to check whether a given switching sequence has the above property. In addition, for the case that $\|S_{i_{k+1}}\cdots S_{i_{k+n}}\|_* = 1 \ \forall k \geq 0, n \geq 1$ for μ-a.e. $i.$, there is no further dichotomy decomposition of $\mathbb{R}^{1\times d}$ into stable manifold E^s and central manifold E^c, except in the case where the support of μ is minimal as in [33, Theorems 2.1 and 2.2] for two-sided Markov shifts and [15, Theorem D′] for $\mathrm{GL}(d, \mathbb{C})$-cocycles driven by one-sided Markov shifts.

Here, we will mainly prove the following more subtle and general results, which implies Theorem B stated in Section 1.2.

Theorem B′. *Assume that S is product bounded; that is equivalent to say, there is a vector norm $\|\cdot\|_*$ on $\mathbb{R}^{1\times d}$, called a "pre-extremal norm" of S, with $\|S_i\|_* \leq 1 \ \forall i \in \mathcal{I}$. Then there hold the following two statements.*

(i) *To every "recurrent switching sequence" $i. = (i_n)_{n=1}^{+\infty} \in \Sigma_{\mathcal{I}}^+$, there corresponds a splitting of $\mathbb{R}^{1\times d}$ into subspaces*

$$\mathbb{R}^{1\times d} = E^s(i.) \oplus E^c(i.),$$

which is independent of the norm $\|\cdot\|_*$, such that

$$\|xS_{i_1}\cdots S_{i_n}\|_2 \to 0 \text{ as } n \to +\infty \quad \forall x \in E^s(i.),$$

and

$$\|xS_{i_1}\cdots S_{i_n}\|_* = \|x\|_* \ \forall n \geq 1 \quad \forall x \in E^c(i.).$$

(ii) *Additionally, let μ be an ergodic measure of θ_+ on $\Sigma_{\mathcal{I}}^+$. Then, for μ-a.e. "weakly Birkhoff-recurrent switching sequence" $i. = (i_n)_{n=1}^{+\infty}$,*

$$\lim_{n\to+\infty} \sqrt[n]{\|x S_{i_1} \cdots S_{i_n}\|_2} < 1 \quad \forall x \in E^s(i.).$$

and

$$\liminf_{n\to+\infty} \|x S_{i_1} \cdots S_{i_n}\|_2 > 0 \quad \forall x \in \mathbb{R}^{1\times d} \setminus E^s(i.).$$

Here the concepts — "recurrent" and "weakly Birkhoff-recurrent" switching sequences — will be precisely defined in Section 4.1 below. Statement **(ii)** of Theorem B′ shows that over almost every weakly Birkhoff-recurrent switching sequences $i.$, for any initial states $x_0 \in \mathbb{R}^{1\times d} \setminus E^s(i.)$, the orbits $(x_n(x_0, i.))_{n=1}^{+\infty}$ of the system S would be far away from the equilibrium $\mathbf{0}_{1\times d}$ as time passes.

Since the closure $\mathrm{Cl}_{\Sigma_{\mathcal{I}}^+}(\{i._{+\ell} \,|\, \ell \geq 0\})$, for any $i. \in W(\theta_+)$ the set of weakly Birkhoff-recurrent switching sequences, is not necessarily a dynamics minimal subset of $\Sigma_{\mathcal{I}}^+$ as shown by [15, Example 5.5], Theorem B′ is an extension of [33, Theorems 2.1 and 2.2] and [15, Theorem D′] from minimal dynamics to non-minimal case.

4.1. Recurrent switching sequences and proof of Theorem B

As is shown in Ian D. Morris [33] and X. Dai [15], the recurrence of a switching sequence is very important for us to study the stability of linear switched systems.

First, we recall from [35, 48] that for a topological dynamical system on a separable metrizable space Υ

$$T \colon \Upsilon \to \Upsilon,$$

a point $y \in \Upsilon$ is called to be "recurrent" by T, provided that one can find a positive integer sequence $n_k \nearrow +\infty$ such that

$$T^{n_k}(y) \to y \quad \text{as } k \to +\infty.$$

In the qualitative theory of ordinary differential equation, this type of recurrent point is also called a "Poisson stable" motion, see, e.g., in [35].

Furthermore, $y \in \Upsilon$ is said to be "weakly Birkhoff-recurrent" by T ([51, 15]), provided that for any $\varepsilon > 0$, there exists an integer $N > 1$ such that

$$\sum_{k=0}^{jN-1} \mathrm{I}_{\mathbb{B}(y,\varepsilon)}(T^k(y)) \geq j \quad \forall j \in \mathbb{N},$$

where $\mathrm{I}_{\mathbb{B}(y,\varepsilon)} \colon \Upsilon \to \{0, 1\}$ is the indicator function of the open ball $\mathbb{B}(y, \varepsilon)$ of radius ε centered at y in the metric space Υ.

From definitions, it is easy to see that a weakly Birkhoff-recurrent point is recurrent (poisson stable), but not vice versa. In particular, for the one-sided Markov shift $\theta_+ \colon \Sigma_{\mathcal{K}}^+ \to \Sigma_{\mathcal{K}}^+$ where $\mathcal{K} = \{1, \ldots, K\}$, a switching sequence $i. = (i_n)_{n=1}^{+\infty} \in \Sigma_{\mathcal{K}}^+$ is recurrent if and only if for any $L \geq 1$ there is an increasing sequence $n_k \nearrow +\infty$ such that

$$(i_{n_k+1}, \ldots, i_{n_k+L}) = (i_1, \ldots, i_L) \quad \forall k \geq 1,$$

for the topology of $\Sigma_{\mathcal{K}}^+$ is generated by the cylinder sets; and $i.$ is weakly Birkhoff-recurrent if and only if it is recurrent and additionally satisfies

$$\limsup_{N \to +\infty} \frac{\mathrm{Card}\{k \colon n_k \le N\}}{N} > 0.$$

We denote by $R(T)$ and $W(T)$, respectively, the sets of all recurrent points and weakly Birkhoff-recurrent points of T. It is easy to see that $R(T)$ and $W(T)$ both are invariant under T such that

$$W(T) \subset R(T).$$

It is easily checked that every periodically switched sequence is weakly Birkhoff-recurrent for the one-sided Markov shift $\theta_+ \colon \Sigma_{\mathcal{I}}^+ \to \Sigma_{\mathcal{I}}^+$. More generally, we have the following basic results.

Poincaré's Recurrence Theorem ([35, Theorem VI3.02], [48, Theorem 1.4])**.** *Let μ be an arbitrary θ_+-invariant probability measure on $\Sigma_{\mathcal{I}}^+$. Then, for μ-a.e. $i. \in \Sigma_{\mathcal{I}}^+$, it is recurrent of θ_+.*

Weak Birkhoff-Recurrence Theorem ([51], [15, Theorem 2.3])**.** *Let μ be an arbitrary θ_+-ergodic probability measure on $\Sigma_{\mathcal{I}}^+$. Then, for μ-a.e. $i. \in \Sigma_{\mathcal{I}}^+$, it is weakly Birkhoff-recurrent of θ_+.*

Here is the important property guaranteed by the recurrence. If $\theta_+^{n_k}(i.) \to i.$ as $k \to +\infty$ for a subsequence $\{n_k\}$, then

$$S_{i_{n_k+1}} \cdots S_{i_{n_k+\ell}} \to S_{i_1} \cdots S_{i_\ell} \quad \text{as } k \to +\infty$$

in $(\mathbb{R}^{d \times d}, \|\cdot\|_2)$, for any $\ell \ge 1$.

Now we are ready to prove Theorem B stated in Section 1.

Proof of Theorem B. From the ergodic theory of Markov chains formulated in Section 2, we only need to consider the associated system $(S, \mu_{p,P})$, driven by the one-sided Markov shift $\theta_+ \colon \Sigma_{\mathcal{K}}^+ \to \Sigma_{\mathcal{K}}^+$. Moreover, from Theorem 2.5, there is no loss of generality in assuming that $\mu_{p,P}$ is ergodic. Then, the statement follows immediately from the Weak Birkhoff-Recurrence Theorem and Theorem B′. $\qquad\square$

As a direct consequence of Theorem B′, we can obtain the following result, which strengthens the statement of [29, Theorem 3].

Corollary 4.1. *Let $S = \{S_1, \dots, S_K\} \subset \mathbb{R}^{d \times d}$ be an arbitrary irreducible set. Then, there always exists at least one θ_+-ergodic probability μ_* on $\Sigma_{\mathcal{K}}^+$ such that there is a splitting of $\mathbb{R}^{1 \times d}$ into weak stable and weak central directions*

$$\mathbb{R}^{1 \times d} = E_{\hat{\rho}}^{ws}(i.) \oplus E_{\hat{\rho}}^{wc}(i.) \quad \text{with } 0 \le \dim E_{\hat{\rho}}^{ws}(i.) \equiv \mathbf{i}_{\mu_*} < d \qquad \text{for } \mu_*\text{-a.e. } i. \in \Sigma_{\mathcal{K}}^+$$

satisfying that

$$\lim_{n \to +\infty} \sqrt[n]{\|x S_{i_1} \cdots S_{i_n}\|_2} < \hat{\rho}(S) \quad \forall x \in E_{\hat{\rho}}^{ws}(i.)$$

and

$$\|xS_{i_1}\cdots S_{i_n}\|_* = \hat{\rho}(S)^n\|x\|_* \quad \forall n \geq 1 \text{ and } x \in E_{\hat{\rho}}^{wc}(i.)$$

for all extremal norms $\|\cdot\|_*$ *of* S. *Here* $\hat{\rho}(S)$ *is the joint spectral radius of* S.

Furthermore, if $\mathbf{i}_{\mu_*} = 0$ *then* S *has the spectral finiteness property; i.e., one can find at least one finite-length word, say* $S_{i_1}\cdots S_{i_m}$, *such that* $\hat{\rho}(S) = \sqrt[m]{\rho(S_{i_1}\cdots S_{i_m})}$.

Proof. Replacing S with $\hat{\rho}(S)^{-1}S$ if necessary, we may assume $\hat{\rho}(S) = 1$ without loss of generality. From [17, Theorem 3.1], it follows that there is at least one ergodic probability μ_* such that

$$1 = \lim_{n\to+\infty} \sqrt[n]{\|S_{i_1}\cdots S_{i_n}\|_2} \quad \text{for } \mu_*\text{-a.e. } i. \in \Sigma_{\mathcal{K}}^+.$$

Since S is irreducible, it is product bounded from [3]. So, to μ_*-a.e. weakly Birkhoff-recurrent switching sequences $i.$, their corresponding splittings $\mathbb{R}^{1\times d} = E^s(i.) \oplus E^c(i.)$ that are given by Theorem B′ satisfy the requirement of Corollary 4.1.

The second part of Corollary 4.1 follows immediately from the classical Gel'fand spectral-radius formula. This completes the proof of Corollary 4.1. □

We note here that since S is irreducible, there always exist extremal norms $\|\cdot\|_*$ of S in Corollary 4.1.

4.2. A basic dichotomy decomposition theorem

Here, in a more general framework, we will prove a basic dichotomous decomposition based on the well-known Poincaré recurrence theorem, by using the idea of "limiting semigroup" due to D.J. Hartfiel [21, 22] and F. Wirth [49]. However, instead of the whole limiting semigroup S^∞ generated by S, we consider only a limiting semigroup over a switching sequence as done in I.D. Morris [33]. Moreover, different with [33] the switching sequence considered here is not necessarily minimal.

Now, our basic decomposition theorem can be stated as follows:

Theorem 4.2. *Let* $T\colon \Upsilon \to \Upsilon$ *be a continuous transformation of a metrizable space* Υ. *Let* $A\colon \Upsilon \to \mathbb{R}^{d\times d}$ *be a continuous family of matrices and* $y \in \Upsilon$ *a recurrent point of* T. *If there is a constant* $\beta_y \geq 1$ *satisfying*

$$\|A_T(n, y)\|_2 \leq \beta_y \quad \forall n \geq 1,$$

then there corresponds a splitting of $\mathbb{R}^{1\times d}$ *into subspaces*

$$\mathbb{R}^{1\times d} = E^s(y) \oplus E^c(y)$$

and a positive integer sequence $n_k \nearrow +\infty$ *with* $T^{n_k}(y) \to y$, *such that*

$$\lim_{k\to+\infty} \|xA_T(n_k, y)\|_2 = 0 \quad \forall x \in E^s(y)$$

and

$$A_T(n_k, y)\restriction E^c(y) \to \mathrm{Id}_{\mathbb{R}^{1\times d}}\restriction E^c(y) \quad as \ k \to +\infty.$$

Here the cocycle $A_T(\cdot, \cdot)$ is defined by

$$A_T(n, y) = A(y) \cdots A(T^{n-1}(y))$$

for any $n \geq 1$ and all $y \in \Upsilon$.

We notice here that for any recurrent point y of T, we cannot guarantee that

$$\liminf_{n \to +\infty} \|x A_T(n, y)\|_2 > 0 \quad \forall x \in \mathbb{R}^{1 \times d} \setminus E^s(y)$$

in the statement of Theorem 4.2, since we are not sure that the decomposition is unique and that there exists a norm $\|\cdot\|_*$ so that $\|A(y)\|_* \leq 1$ for all $y \in \Upsilon$ except the case driven by the one-sided or two-sided Markov shift.

Proof. For any recurrent point y of T, put

$$\mathcal{S}(y) = \left\{ B \in \mathbb{R}^{d \times d} \mid \exists n_k \nearrow +\infty \text{ s.t. } T^{n_k}(y) \to y \text{ and } \lim_{k \to +\infty} A_T(n_k, y) = B \right\}.$$

Since the cocycle $A_T(\cdot, y)$ is uniformly bounded and the point y is T-recurrent, $\mathcal{S}(y)$ is not empty and bounded in $\mathbb{R}^{d \times d}$. We further claim that it is a compact semigroup in the sense of matrix multiplication.

Firstly, let $B = \lim_{\ell \to +\infty} B_\ell$ where $B_\ell \in \mathcal{S}(y)$. Then for any $\ell \geq 1$, one can choose $n_{k_\ell} \geq \ell$ satisfying

$$\text{dist}\,(T^{n_{k_\ell}}(y), y) \leq \frac{1}{2\ell} \quad \text{and} \quad \|A_T(n_{k_\ell}, y) - B_\ell\|_2 \leq \frac{1}{2\ell}.$$

So, $B \in \mathcal{S}(y)$. This implies that $\mathcal{S}(y)$ is closed and hence compact in the space $(\mathbb{R}^{d \times d}, \|\cdot\|_2)$.

Secondly, let $B_1, B_2 \in \mathcal{S}(y)$. Then one can find two sequences $n_k^{(1)} \nearrow +\infty$ and $n_k^{(2)} \nearrow +\infty$ such that

$$T^{n_k^{(1)}}(y) \to y, \; A_T(n_k^{(1)}, y) \to B_1 \quad \text{and} \quad T^{n_k^{(2)}}(y) \to y, \; A_T(n_k^{(2)}, y) \to B_2 \qquad \text{as } k \to +\infty.$$

To show $B_2 B_1 \in \mathcal{S}(y)$, it suffices to prove that for any $N \geq 1$ and $\varepsilon > 0$, there is some $n > N$ such that

$$\text{dist}\,(T^n(y), y) < \varepsilon \quad \text{and} \quad \|A_T(n, y) - B_2 B_1\|_2 < \varepsilon.$$

In fact, there first exists some $k \geq N$ such that

$$\text{dist}\left(T^{n_k^{(1)}}(y), y\right) < \frac{\varepsilon}{2} \quad \text{and} \quad \|B_2 A_T(n_k^{(1)}, y) - B_2 B_1\|_2 < \frac{\varepsilon}{3}.$$

Then for the taken k, there is some $K' \geq N$ such that for any $k' \geq K'$,

$$\|A_T(n_{k'}^{(2)}, y) A_T(n_k^{(1)}, y) - B_2 A_T(n_k^{(1)}, y)\|_2 < \frac{\varepsilon}{3}.$$

Secondly, since $T^{n_{k'}^{(2)}}(y)$ converges to y as $k' \to +\infty$ and $A_T(\cdot, y)$ is uniformly bounded, from the continuity of $A_T(n, \cdot)$ one can find some $k' > K'$ such that

$$\text{dist}\left(T^{n_k^{(1)}}(T^{n_{k'}^{(2)}}(y)), T^{n_k^{(1)}}(y)\right) < \frac{\varepsilon}{2}$$

and

$$\|A_T(n_{k'}^{(2)}, y)A_T(n_k^{(1)}, T^{n_{k'}^{(2)}}(y)) - A_T(n_{k'}^{(2)}, y)A_T(n_k^{(1)}, y)\|_2 < \frac{\varepsilon}{3}.$$

Because for $n = n_k^{(1)} + n_{k'}^{(2)}$ we have

$$\|A_T(n, y) - B_2B_1\|_2 \le \|B_2A_T(n_k^{(1)}, y) - B_2B_1\|_2 + \|A_T(n_{k'}^{(2)}, y)A_T(n_k^{(1)}, y) - B_2A_T(n_k^{(1)}, y)\|_2$$
$$+ \|A_T(n_{k'}^{(2)}, y)A_T(n_k^{(1)}, T^{n_{k'}^{(2)}}(y)) - A_T(n_{k'}^{(2)}, y)A_T(n_k^{(1)}, y)\|_2$$
$$< \varepsilon$$

and

$$\text{dist}\,(T^n(y), y) \le \text{dist}\left(T^{n_k^{(1)}}(T^{n_{k'}^{(2)}}(y)), T^{n_k^{(1)}}(y)\right) + \text{dist}\left(T^{n_k^{(1)}}(y), y\right) < \varepsilon$$

we obtain that $B_2B_1 \in \mathcal{S}(y)$ and similarly $B_1B_2 \in \mathcal{S}(y)$. Thus, $\mathcal{S}(y)$ is a compact semigroup.

Then, $\mathcal{S}(y)$ contains an idempotent element P, i.e., $P^2 = P$ (see, e.g. [24]). Next we define

$$E^s(y) = \text{the kernel of } P(\cdot) \quad \text{and} \quad E^c(y) = \text{the range of } P(\cdot),$$

where $P(\cdot)\colon \mathbb{R}^{1 \times d} \to \mathbb{R}^{1 \times d}$ is defined by $x \mapsto xP$. Then

$$\mathbb{R}^{1 \times d} = E^s(y) \oplus E^c(y).$$

Since P is an ω-limit point of the sequence $\{A_T(n, y)\}_{n=1}^{+\infty}$, there is a subsequence $\{n_k\}_{k=1}^{+\infty}$ such that $A_T(n_k, y) \to P$ as $k \to +\infty$. This implies that

$$\lim_{k \to +\infty} xA_T(n_k, y) = \mathbf{0}_{1 \times d} \ \forall x \in E^s(y) \quad \text{and} \quad \lim_{k \to +\infty} xA_T(n_k, y) = x \ \forall x \in E^c(y).$$

This thus proves the statement of Theorem 4.2. □

A special case of Theorem 4.2 is the following statement.

Corollary 4.3. *Let* $T\colon \Upsilon \to \Upsilon$ *be a continuous transformation of a metrizable space* Υ *having a recurrent point y. Let* $A\colon \Upsilon \to \mathbb{R}^{d \times d}$ *be a continuous family of matrices satisfying*

$$\|A_T(n, y)\|_2 \le \beta_y \quad \forall n \ge 1, \qquad \text{for some constant } \beta_y \ge 1.$$

If $E^s(y) = \{\mathbf{0}_{1 \times d}\}$, *then*

$$A_T(n_k, y) \to \text{Id}_{\mathbb{R}^{1 \times d}} \quad \text{as } k \to +\infty$$

for some sequence $n_k \to +\infty$.

We have the following two remarks about the important Theorem 4.2:

Remark 4.4. Theorem 4.2 may not be true if $i.\colon \mathbb{N} \to \mathcal{I}$ is not a recurrent switching sequence of S. For example, let $S = \{S_0, S_1\}$, where $\|S_0\|_2 < 1$ and $S_1 = \text{Id}_{\mathbb{R}^{d \times d}}$ is the identity matrix. Clearly, S is product bounded. However, for the non-recurrent switching sequence $i.\colon \mathbb{N} \to \{0, 1\}$ where $i_1 = 0$ and $i_n = 1$ for all $n \ge 2$, we have

$$\lim_{n \to +\infty} xS_{i_1} \cdots S_{i_n} = xS_0 \quad \forall x \in \mathbb{R}^{1 \times d}.$$

which implies that the statement of Theorem 4.2 is not true if $S_0 \ne \mathbf{0}_{d \times d}$.

Remark 4.5. It is natural to ask whether the convergence is exponentially fast over the stable manifold $E^s(i.)$; that is,

$$\limsup_{n \to +\infty} \frac{1}{n} \log \|x S_{i_1} \cdots S_{i_n}\|_2 < 0, \qquad \forall x \in E^s(i.).$$

The following example shows this may not be true.

Example 4.6. Let $S = \{S_0, S_1\}$ with

$$S_0 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad \text{and} \quad S_1 = \begin{bmatrix} \alpha & 0 \\ 0 & 1 \end{bmatrix}$$

where $0 < \alpha < 1$. We now construct a recurrent switching sequence by induction. For a word $w = (i_1 \cdots i_k) \in \{0, 1\}^k$, let $|w| = k$ denote the length of the word w and O_k stand for the word consisting of k numbers of 0, i.e, $O_k = (0, \ldots, 0) \in \{0, 1\}^k$. For any pair $w_1 = (i_1, \ldots, i_k)$ and $w_2 = (i'_1, \ldots, i'_m)$, put $w_1 w_2 = (i_1, \ldots, i_k, i'_1, \ldots, i'_m) \in \{0, 1\}^{k+m}$
   Let $\sigma_1 = (1)$, $\sigma_2 = (\sigma_1 O_{|\sigma_1|^2} \sigma_1)$. Inductively, for $n \geq 2$, let

$$\sigma_n = (\sigma_{n-1} O_{|\sigma_{n-1}|^2} \sigma_{n-1}) \quad \text{and} \quad i. = \lim_{n \to +\infty} \sigma_n.$$

It is easy to see from the construction that $i.$ is a recurrent switching sequence of S and

$$\mathbb{R}^{1 \times 2} = E^s(i.) \oplus E^c(i.) \quad \text{where } E^s(i.) = \{(x_1, 0) \mid x_1 \in \mathbb{R}\} \text{ and } E^c(i.) = \{(0, x_2) \mid x_2 \in \mathbb{R}\}.$$

A routine check shows that

$$\limsup_{n \to +\infty} \frac{1}{n} \log \|x S_{i_1} \cdots S_{i_n}\|_2 = \lim_{n \to +\infty} \frac{n}{n^2} \log \alpha = 0 \quad \forall x \in E^s(i.) \setminus \{\mathbf{0}_{1 \times 2}\}.$$

So the convergence is not exponentially fast over the switching sequence $i.$.

   In Section 4.3 below, we will further consider the regularity of the stable manifold $E^s(y)$ and prove that restricted to it, $\|A_T(n, y) \restriction E^s(y)\|_2$ converges exponentially fast to 0 as $n \to +\infty$ if, in addition, y is a weakly Birkhoff-recurrent switching sequence.

4.3. The measurability of the stable manifolds
   For $0 \leq p \leq d$, we let $\mathscr{G}(p, \mathbb{R}^{1 \times d})$ denote the set of all p-dimensional subspaces of $\mathbb{R}^{1 \times d}$ and $\mathscr{G}(\mathbb{R}^{1 \times d}) = \bigsqcup_{p=0}^{d} \mathscr{G}(p, \mathbb{R}^{1 \times d})$, where $\bigsqcup$ means the disjoint union. We equip $\mathscr{G}(\mathbb{R}^{1 \times d})$ with the compact topology induced by the Hausdorff metric $d_H(\cdot, \cdot)$, i.e., for any $V, W \in \mathscr{G}(\mathbb{R}^{1 \times d})$,

$$d_H(V, W) = \max \left\{ \sup_{v \in V^\sharp} \inf_{w \in W^\sharp} \|v - w\|_2, \sup_{w \in W^\sharp} \inf_{v \in V^\sharp} \|w - v\|_2 \right\} \quad \text{where } V^\sharp = \{v \in V : \|v\|_2 = 1\}.$$

Here $\| \cdot \|_2$ is the euclidean vector-norm on $\mathbb{R}^{1 \times d}$ as before.
   Let $S \colon \mathcal{I} \ni i \mapsto S_i \in \mathbb{R}^{d \times d}$ be a continuous matrix-valued mapping defined on a topological space $\mathcal{I}$. For any $i. = (i_n)_{n=1}^{+\infty} \in \Sigma_{\mathcal{I}}^+$, we set

$$E^s(i.) = \left\{ x \in \mathbb{R}^{1 \times d} \mid x S_{i_1} \cdots S_{i_n} \to \mathbf{0}_{1 \times d} \text{ as } n \to +\infty \right\}.$$

It is a subspace of $\mathbb{R}^{1\times d}$ with $\mathrm{S}_{i_1}(E^s(i.)) \subseteq E^s(i._{+1})$ and called the "stable manifold/direction" of S over the switching sequence $i. = (i_n)_{n=1}^{+\infty}$. Now, we will consider the measurability of the following functions

$$\Sigma_{\mathcal{I}}^+ \ni i. \mapsto E^s(i.) \in \mathscr{G}(\mathbb{R}^{1\times d}) \quad \text{and} \quad \Sigma_{\mathcal{I}}^+ \ni i. \mapsto \dim E^s(i.) \in \mathbb{Z}_+.$$

For that, we need the following lemma, which is a direct corollary of the classical Oseledec multiplicative ergodic theorem (see [37] and also [48, Theorem 10.2]).

Lemma 4.7 (Oseledec). *Let $T\colon (\Omega, \mathscr{F}, \mathbb{P}) \to (\Omega, \mathscr{F}, \mathbb{P})$ be an ergodic measure-preserving transformation of a probability space $(\Omega, \mathscr{F}, \mathbb{P})$. Let $A\colon \Omega \to \mathbb{R}^{d\times d}$ be a measurable family of matrices satisfying*

$$\log^+ \|A(\cdot)\|_2 \in L^1(\Omega, \mathscr{F}, \mathbb{P}), \quad \text{where } x^+ = \max\{0, x\}.$$

Then, there exists $\Omega' \in \mathscr{F}$ with $T(\Omega') \subseteq \Omega'$ and $\mathbb{P}(\Omega') = 1$ having the following properties.

(1) *There exists a measurable function $\omega \mapsto V^s(\omega) \in \mathscr{G}(\mathbb{R}^{1\times d})$ and an integer $0 \le \mathbf{i}_{\mathbb{P}} \le d$ with $\dim V^s(\omega) \equiv \mathbf{i}_{\mathbb{P}}$ for all $\omega \in \Omega'$, i.e., $V^s(\omega) \in \mathscr{G}(\mathbf{i}_{\mathbb{P}}, \mathbb{R}^{1\times d})$.*

(2) *If $\omega \in \Omega'$ then*

$$\chi(\omega, v) = \lim_{n\to+\infty} \frac{1}{n} \log \|vA_T(n, \omega)\|_2 < 0 \qquad \forall v \in V^s(\omega)$$

and

$$\chi(\omega, v) = \lim_{n\to+\infty} \frac{1}{n} \log \|vA_T(n, \omega)\|_2 \ge 0 \qquad \forall v \in \mathbb{R}^{1\times d} \setminus V^s(\omega).$$

Here the cocycle

$$A_T(n, \omega) = A(\omega) \cdots A(T^{n-1}\omega)$$

for any $n \ge 1$ and all $\omega \in \Omega$.

If $E^s(i.) = V^s(i.)$ a.e. for S, then the desired measurability holds. From the following theorem, this is the case under an additional condition — product boundedness.

Theorem 4.8. *Let $T\colon (\Omega, \mathscr{F}, \mathbb{P}) \to (\Omega, \mathscr{F}, \mathbb{P})$ be an ergodic measure-preserving continuous transformation of a Borel probability space $(\Omega, \mathscr{F}, \mathbb{P})$ based on a separable metrizable space Ω. Let $A\colon \Omega \to \mathbb{R}^{d\times d}$ be a continuous family of matrices satisfying*

$$\|A_T(n, \omega)\|_2 \le \beta \quad \forall n \ge 1 \text{ and } \omega \in \Omega, \qquad \text{for some constant } \beta \ge 1.$$

Then there exists $\Omega' \in \mathscr{F}$ with $T(\Omega') \subseteq \Omega'$ and $\mathbb{P}(\Omega') = 1$ having the following properties.

(1) *There exists a measurable function $\omega \mapsto V^s(\omega) \in \mathscr{G}(\mathbb{R}^{1\times d})$ and an integer $0 \le \mathbf{i}_{\mathbb{P}} \le d$ with $\dim V^s(\omega) \equiv \mathbf{i}_{\mathbb{P}}$ for all $\omega \in \Omega'$.*

(2) *If $\omega \in \Omega'$ then*

$$\chi(\omega, v) = \lim_{n \to +\infty} \frac{1}{n} \log \|vA_T(n, \omega)\|_2 < 0 \qquad v \in V^s(\omega)$$

and

$$\chi(\omega, v) = \lim_{n \to +\infty} \frac{1}{n} \log \|vA_T(n, \omega)\|_2 = 0 \qquad v \in \mathbb{R}^{1 \times d} \setminus V^s(\omega).$$

(3) *If ω belongs to Ω', then*

$$\liminf_{n \to +\infty} \|vA_T(n, \omega)\|_2 > 0 \qquad \forall v \in \mathbb{R}^{1 \times d} \setminus V^s(\omega).$$

This theorem will be proven later. We first note here that if $S = \{S_i\}_{i \in \mathcal{I}} \subset \mathbb{R}^{d \times d}$ is product bounded then from, e.g., [11, 28], there exists a pre-extremal norm $\| \cdot \|_*$ on $\mathbb{R}^{1 \times d}$, i.e., $\|S_i\|_* \le 1$ for all $i \in \mathcal{I}$. This might simplify many arguments as done in [15]. However, in the framework of the above Theorem 4.8, there is no such a pre-extremal norm on $\mathbb{R}^{1 \times d}$, for the cocycle $A_T(\cdot, \cdot)$.

Then it is time to prove Theorem B$'$.

Proof of Theorem B$'$. Applying Theorem 4.8 to the situation of Theorem B$'$, we see that

$$E^s(i.) = V^s(i.) \quad \mu\text{-a.e. } i. \in R(\theta_+),$$

where $R(\theta_+)$ is the recurrent point set. Therefore, combining the Weak Birkhoff-Recurrence Theorem, Theorems 4.2 and 4.8 completes the proof of Theorem B$'$. $\qquad \square$

Next, we will devote our attention to proving Theorem 4.8. To prove property **(3)** of Theorem 4.8, we will need the following simple result.

Lemma 4.9. *Under the situation of Theorem 4.8, for any subspace $L \subseteq \mathbb{R}^{1 \times d}$ and any $\omega \in \Omega$, the following statements are equivalent to each other:*

(a) $\lim_{n \to +\infty} \|A_T(n, \omega) \restriction L\|_2 = 0$.

(b) $\liminf_{n \to +\infty} \|A_T(n, \omega) \restriction L\|_2 = 0$.

(c) $\liminf_{n \to +\infty} \|vA_T(n, \omega)\|_2 = 0 \; \forall v \in L$.

Proof. We only need to prove **(c)** $\Rightarrow$ **(a)**. Let $\varepsilon > 0$ be arbitrary and then take $\delta > 0$ so small that satisfying $\max\{\delta\beta, \delta\beta^2\} < \varepsilon$, where β is given by product boundedness of A_T as in Theorem 4.8. As $L^\sharp = \{v \in L: \|v\|_2 = 1\}$ is compact and $vA_T(n, \omega))$ is continuous with respect to $v \in L^\sharp$, it follows, from **(c)**, that one can find some integer $N \ge 1$ such that $\|vA_T(N, \omega)\|_2 \le \delta\beta \; \forall v \in L^\sharp$. Then for any $n \ge N$, we have $\|A_T(n, \omega)\restriction L\|_2 \le \delta\beta^2 < \varepsilon$.

This completes the proof of Lemma 4.9. $\qquad \square$

We now continue the proof of Theorem 4.8.

Proof of Theorem 4.8. First it easily follows, from Lemma 4.7, that there exists $\Omega' \in \mathscr{F}$ with $T(\Omega') \subseteq \Omega'$ and $\mathbb{P}(\Omega') = 1$ having properties **(1)** and **(2)** of Theorem 4.8.

For any $\omega \in \Omega$, define the subspace

$$E_{A,T}^s(\omega) = \left\{ x \in \mathbb{R}^{1 \times d} : \lim_{n \to +\infty} \|xA_T(n, \omega)\|_2 = 0 \right\}.$$

From Lemma 4.9 with $L = \mathrm{span}\{x\}$, $x \in E_{A,T}^s(\omega)$ if and only if $\liminf_{n \to +\infty} \|xA_T(n, \omega)\|_2 = 0$. It is easily checked that for all $\omega \in \Omega'$,

$$\lim_{n \to +\infty} \|A_T(n, \omega) \upharpoonright E_{A,T}^s(\omega)\|_2 = 0, \quad A(\omega)(E_{A,T}^s(\omega)) \subseteq E_{A,T}^s(T(\omega)) \quad \text{and} \quad V^s(\omega) \subseteq E_{A,T}^s(\omega).$$

Next, we will prove that $\dim E_{A,T}^s(\omega) \equiv$ constant for a.e. $\omega \in \Omega$. At first, for any $\omega \in \Omega'$, since $xA(\omega) = \mathbf{0}_{1 \times d}$ implies $x \in V^s(\omega)$, the following lemma comes immediately from Lemma 4.7.

Lemma 4.10. *Under the situation of Theorem 4.8, for any $\omega \in \Omega'$, it holds that*

$$\dim E_{A,T}^s(T^n(\omega)) = \dim E_{A,T}^s(\omega)$$

for each $n \geq 1$.

For the Borel probability space $(\Omega, \mathscr{F}, \mathbb{P})$, let $\mathscr{F}^{\mathbb{P}}$ be the "$\mathbb{P}$-completion" of $\mathscr{F}$, as the σ-field $\mathscr{F}^{\mathbb{P}} = \sigma(\mathscr{F} \cup \mathscr{N}_{\mathbb{P}})$, where $\mathscr{N}_{\mathbb{P}}$ denotes the class of all subsets of arbitrary $\mathbb{P}$-null sets in $\mathscr{F}$. Then $\mathbb{P}$ on $\mathscr{F}$ has a unique extension to the σ-field $\mathscr{F}^{\mathbb{P}}$ and $T: (\Omega, \mathscr{F}^{\mathbb{P}}, \mathbb{P}) \to (\Omega, \mathscr{F}^{\mathbb{P}}, \mathbb{P})$ is also an ergodic measure-preserving transformation. This completion enables us using some classical results in measure theory.

Lemma 4.11. *Let $(Y, \mathscr{F}_Y)$ be a Borel measurable space. Then a function*

$$f: \Omega \to Y$$

is $\mathscr{F}^{\mathbb{P}}$-measurable if and only if there exists some $\mathscr{F}$-measurable function $g: \Omega \to Y$ satisfying $f(\omega) = g(\omega)$ for $\mathbb{P}$-a.e. $\omega \in \Omega$.

This result is well known and can be found in many textbooks on real analysis and probability, for example, in P. Billinsley [8]. The following is an other classical result needed.

Lemma 4.12 (Projection and Sections, Lusin, Choquet, Meyer; see [27, Theorem A1.4])**.** *Let $(Y, \mathscr{F}_Y)$ be a Borel measurable space and $\pi: \Omega \times Y \to \Omega$ the projection defined by $(\omega, y) \mapsto \omega$. Then for any $B \in \mathscr{F} \otimes \mathscr{F}_Y$,*

(i) *$\pi(B)$ belongs to $\mathscr{F}^{\mathbb{P}}$;*

(ii) *there exists a $\mathscr{F}^{\mathbb{P}}$-measurable element $\eta: \Omega \to Y$ such that $(\omega, \eta(\omega)) \in B$ for $\mathbb{P}$-a.e. on $\pi(B)$.*

For any $1 \leq p \leq d$, let

$$g_p: \Omega \times \mathscr{G}(p, \mathbb{R}^{1 \times d}) \to [0, \infty); \quad (\omega, L) \mapsto \liminf_{n \to +\infty} \|A_T(n, \omega) \upharpoonright L\|_2$$

and

$$\pi_p: \Omega \times \mathscr{G}(p, \mathbb{R}^{1 \times d}) \to \Omega; \quad (\omega, L) \mapsto \omega.$$

Moreover, set

$$\Omega_p = \Omega' \cap \pi_p(g_p^{-1}\{0\}).$$

Although g_p is Borel measurable and so the pre-image $g_p^{-1}\{0\}$ is a Borel subset of the product space $\Omega \times \mathscr{G}(p, \mathbb{R}^{1 \times d})$, yet the projection Ω_p need not belong to the Borel σ-field $\mathscr{F}$ of Ω.

Lemma 4.13. *Under the situation of Theorem 4.8, for $\mathbf{i}_{\mathbb{P}} \le p \le d$, Ω_p is such that $T(\Omega_p) \subseteq \Omega_p$ and $\Omega_p \in \mathscr{F}^{\mathbb{P}}$, where $\mathbf{i}_{\mathbb{P}}$ is the stable index given by Lemma 4.7.*

Proof of Lemma 4.13. Let $\mathbf{i}_{\mathbb{P}} \le p \le d$. We note that for any $\omega \in \Omega_p$, from Lemma 4.7 one can always find some $L \in \mathscr{G}(p, \mathbb{R}^{1 \times d})$ with $V^s(\omega) \subseteq L$ such that $(\omega, L) \in g_p^{-1}\{0\}$. Then

$$\dim \left(V^s(T(\omega)) + A(\omega)(L) \right) \ge p,$$

which implies that there is some $M \in \mathscr{G}(p, \mathbb{R}^{1 \times d})$ with $M \subseteq V^s(T(\omega)) + A(\omega)(L)$. This means that $(T(\omega), M) \in g_p^{-1}\{0\}$ and hence $T(\omega) \in \Omega_p$. So, the invariance holds. The measurability comes from Lemma 4.12. This proves Lemma 4.13. $\qquad\square$

Since $\mathbb{P}$ is T-ergodic, $\mathbb{P}(\Omega_p) = 1$ or 0 for any $\mathbf{i}_{\mathbb{P}} \le p \le d$. It is easily checked that $\mathbb{P}(\Omega_{\mathbf{i}_{\mathbb{P}}}) = 1$. Let $\mathbf{i}'_{\mathbb{P}}$ be the maximal integer p such that $\mathbb{P}(\Omega_p) = 1$. Then from Lemma 4.12, we could obtain the following

Lemma 4.14. *Under the situation of Theorem 4.8, one can find a $\mathscr{F}^{\mathbb{P}}$-measurable element*

$$\eta \colon \Omega \to \mathscr{G}(\mathbf{i}'_{\mathbb{P}}, \mathbb{R}^d)$$

such that

$$(\omega, \eta(\omega)) \in g_{\mathbf{i}'_{\mathbb{P}}}^{-1}\{0\} \quad and \quad \eta(\omega) = E^s_{A,T}(\omega)$$

for $\mathbb{P}$-a.e. $\omega \in \Omega_{\mathbf{i}'_{\mathbb{P}}}$.

This, combining with Lemma 4.11, implies that there exists a $\mathscr{F}$-measurable element, written as

$$\boldsymbol{\xi} \colon \Omega \to \mathscr{G}(\mathbf{i}'_{\mathbb{P}}, \mathbb{R}^d),$$

such that

$$\boldsymbol{\xi}(\omega) = E^s_{A,T}(\omega) \quad \text{for } \mathbb{P}\text{-a.e. } \omega \in \Omega.$$

Clearly, $\mathbf{i}_{\mathbb{P}} \le \mathbf{i}'_{\mathbb{P}}$. To complete the proof of Theorem 4.8, it is sufficient to prove $\mathbf{i}_{\mathbb{P}} = \mathbf{i}'_{\mathbb{P}}$. This case comes immediately from the following.

Lemma 4.15. *Under the situation of Theorem 4.8, the following limits exist and*

$$\lim_{n \to +\infty} \frac{1}{n} \log \|A_T(n, \omega) \restriction E^s_{A,T}(\omega)\|_2 < 0$$

for $\mathbb{P}$-a.e. $\omega \in \Omega$.

Proof of Lemma 4.15. As the section $\boldsymbol{\xi}$, defined above, is Borel measurable, from Lusin's theorem it follows that there is a closed subset $E \subset \Omega'$ with $\mathbb{P}(E) > 0$ such that

$$\omega \mapsto \boldsymbol{\xi}(\omega) = E_{A,T}^s(\omega)$$

is continuous on E. Moreover, from the Birkhoff ergodic theorem, it follows that there is a T-invariant Borel subset $\Omega'' \subseteq \Omega'$ with $\mathbb{P}(\Omega'') = 1$ such that

$$\boldsymbol{\xi}(\omega) = E_{A,T}^s(\omega) \quad \text{and} \quad \lim_{n \to +\infty} \frac{1}{n} \sum_{i=0}^{n-1} I_E(T^i(\omega)) = \mathbb{P}(E)$$

for any $\omega \in \Omega''$, where $I_E(\cdot)$ is the indicator of E defined on Ω.

Let $0 < \varepsilon \ll 1$ be an arbitrary constant. Similar to the proof of Lemma 4.9, one can take, by the continuity of $\boldsymbol{\xi}$ restricted to E, some integer $N \geq 1$ sufficiently large such that

$$\|A_T(N, \omega) \upharpoonright E_{A,T}^s(\omega)\|_2 < \varepsilon \quad \forall \omega \in E.$$

In what follows, let $\hat{\omega} \in \Omega''$ be arbitrarily given. Then, there is some integer $\ell_{\hat{\omega}} \gg N$ such that

$$\sum_{i=0}^{j\ell_{\hat{\omega}}-1} I_E(T^i(\hat{\omega})) \geq j \qquad \forall j \geq 1.$$

This implies that for any $j \geq 1$ one could find j integers, say $\check{k}_1, \ldots, \check{k}_j$, such that

$$0 \leq \check{k}_1 < \check{k}_2 < \cdots < \check{k}_j \leq j\ell_{\hat{\omega}} - 1 \quad \text{and} \quad T^{\check{k}_s}(\hat{\omega}) \in E \ \text{ for } 1 \leq s \leq j.$$

As j is big sufficiently, we can choose at least $[j/N]$ integers, say $\check{k}_{j_1}, \ldots, \check{k}_{j_{[j/N]}}$, from $\{\check{k}_1, \ldots, \check{k}_j\}$, such that

$$0 \leq \check{k}_{j_1} < \check{k}_{j_2} < \cdots < \check{k}_{j_{[j/N]}} \leq j\ell_{\hat{\omega}} - 1$$

$$\check{k}_{j_{s-1}} + N \leq \check{k}_{j_s} \quad \text{for } 2 \leq s \leq [j/N]$$

and

$$T^{\check{k}_{j_s}}(\hat{\omega}) \in E \quad \text{for } 1 \leq s \leq [j/N],$$

where $[j/N] = \max\{n \in \mathbb{Z} \,|\, n \leq j/N\}$. Therefore,

$$\begin{aligned}
\limsup_{n \to +\infty} \frac{1}{n} \log \|A_T(n, \hat{\omega}) \upharpoonright E_{A,T}^s(\hat{\omega})\|_2 &= \limsup_{j \to +\infty} \frac{1}{j\ell_{\hat{\omega}}} \log \|A_T(j\ell_{\hat{\omega}}, \hat{\omega}) \upharpoonright E_{A,T}^s(\hat{\omega})\|_2 \\
&= \limsup_{j \to +\infty} \frac{1}{j\ell_{\hat{\omega}}} \log \|A_T(j\ell_{\hat{\omega}} + N, \hat{\omega}) \upharpoonright E_{A,T}^s(\hat{\omega})\|_2 \\
&\leq \limsup_{j \to +\infty} \frac{1}{j\ell_{\hat{\omega}}} \left\{ \log \varepsilon^{[j/N]} + \log \beta^{1+[j/N]} \right\} \\
&= \frac{1}{N\ell_{\hat{\omega}}} \log(\varepsilon\beta) \\
&< 0
\end{aligned}$$

because $\varepsilon > 0$ is arbitrary. Hence, $E_{A,T}^s(\hat{\omega}) = V^s(\hat{\omega})$ for any $\hat{\omega} \in \Omega''$.

This proves Lemma 4.15. $\qquad\square$

Now, the statement of Theorem 4.8 follows immediately from Lemmas 4.7, 4.9 and 4.15. Thus, the proof of Theorem 4.8 is completed. $\qquad\square$

4.4. A remark to Theorem B′

Why do we only consider μ-a.e. $i. \in W(\theta_+)$, not every $i. \in W(\theta_+)$, in statement **(ii)** of Theorem B′? The reason is this: although the recurrence of a weakly Birkhoff-recurrent switching sequence $i. = (i_n)_{n=1}^{+\infty}$ is so strong that having a positive recurrent frequency, yet the stable manifold $E^s(i._{+n_\ell})$ need not approximate sufficiently $E^s(i.)$ even $i._{+n_\ell}$ converges to $i.$ as $\ell \to +\infty$, because the splitting $E^s(i.) \oplus E^c(i.)$ of $\mathbb{R}^{1 \times d}$ defined by statement **(i)** of Theorem B′ is not necessarily continuous with respect to $i. \in R(\theta_+)$. This is just one of the essential hard points of nonuniformly hyperbolic systems.

In light of this reason, we have to apply the Oseledec multiplicative ergodic theorem to obtain a weak regularity — the measurability of $E^s(i.)$, with respect to μ-a.e. $i. \in \Sigma_{\mathcal{I}}^+$, as done in Section 4.3.

However, if $E^s(i.) = \mathbb{R}^{1 \times d}$, then $E^s(i._{+\ell}) = \mathbb{R}^{1 \times d}$ for all $\ell \geq 1$. And the above short point is naturally avoided in this case. So, as done in the proof of Lemma 4.15 that is independently interesting, together with Theorem 4.2, we can easily obtain the following, which is the counterpart of Corollary 4.3.

Corollary 4.16 ([15, Theorem 2.4]). *Let* $T \colon \Omega \to \Omega$ *be a continuous transformation of a separable metrizable space* Ω. *Let* $A \colon \Omega \to \mathbb{R}^{d \times d}$ *be a continuous family of matrices satisfying*

$$\|A_T(n, \omega)\|_2 \leq \beta \quad \forall\, n \geq 1 \text{ and } \omega \in \Omega \quad \text{for some constant } \beta \geq 1.$$

If $\omega \in W(T)$ *satisfies* $E_{A,T}^s(\omega) = \mathbb{R}^{1 \times d}$, *then* $\|A_T(n, \omega)\|_2$ *converges exponentially fast to* 0 *as* $n \to +\infty$.

Finally, for the convenience of our subsequent papers, we reformulate Theorem B′ as follows:

Theorem B″. *Let* $T \colon (\Omega, \mathscr{F}, \mathbb{P}) \to (\Omega, \mathscr{F}, \mathbb{P})$ *be an ergodic measure-preserving continuous transformation of a Borel probability space* $(\Omega, \mathscr{F}, \mathbb{P})$ *based on a separable metrizable space* Ω. *Let* $A \colon \Omega \to \mathbb{R}^{d \times d}$ *be a continuous family of matrices satisfying*

$$\|A_T(n, \omega)\|_2 \leq \beta \quad \forall n \geq 1 \text{ and } \omega \in \Omega, \qquad \text{for some constant } \beta \geq 1.$$

Then there hold the following two statements.

(1) *To every recurrent point* $\omega \in R(T)$, *there corresponds a splitting of* $\mathbb{R}^{1 \times d}$ *into subspaces*

$$\mathbb{R}^{1 \times d} = E^s(\omega) \oplus E^c(\omega),$$

and a positive integer sequence $n_k \nearrow +\infty$ *with* $T^{n_k}(\omega) \to \omega$, *such that*

$$\lim_{k \to +\infty} \|A_T(n_k, \omega) \restriction E^s(\omega)\|_2 = 0$$

and

$$A_T(n_k, \omega) \restriction E^c(\omega) \to \mathrm{Id}_{\mathbb{R}^{1 \times d}} \restriction E^c(\omega) \quad \text{as } k \to +\infty.$$

(2) *For* $\mathbb{P}$-a.e. weakly Birkhoff-recurrent point $\omega \in W(T)$,

$$\lim_{n \to +\infty} \sqrt[n]{\|xA_T(n, \omega)\|_2} < 1 \ \forall x \in E^s(\omega) \quad \text{and} \quad \liminf_{n \to +\infty} \|xA_T(n, \omega)\|_2 > 0 \ \forall x \in \mathbb{R}^{1 \times d} \setminus E^s(\omega).$$

5. Concluding remarks

For a Markovian jump linear system, we introduced two concepts — pointwise convergence and pointwise exponential convergence. The latter is expected in many aspects, like numerical computation, optimization control and so on. These two kinds of stabilizabilities are not equivalent to each other, in general. However, we showed that if the Markovian jump linear system is product bounded, then the pointwise convergence and pointwise exponential convergence are equivalent to each other.

Acknowledgments

References

[1] A. BACCIOTTI AND R. LIONEL, *Regularity of Liapunov functions for stable systems*, System & Control Letters, 41 (2000), pp. 265–270.

[2] A. BACCIOTTI AND R. LIONEL, *Liapunov Functions and Stability in Control Theory*, 2nd ed., Comm. Control Engrg. Ser., Springer-Verlag, Berlin 2005.

[3] N. BARABANOV, *Lyapunov indicators of discrete inclusions I–III*, Autom. Remote Control, 49 (1988), pp. 152–157, 283–287, 558–565.

[4] N. BARABANOV, *Absolute characteristic exponent of a class of linear non-stationary systems of differential equations*, Siberian Math. J., 29 (1988), pp. 521–530.

[5] N. BARABANOV, *On the Aizerman problem for 3rd-order nonstationary systems*, Differ. Equ., 29 (1993), pp. 1439–448.

[6] G.P. BARKER, L.T. CONNER, AND D.P. STANFORD, *Complete controllability and contractability in multi-modal systems*, Linear Algebra Appl., 110 (1988), pp. 55–74.

[7] M.A. BERGER AND Y. WANG, *Bounded semigroups of matrices*, Linear Algebra Appl., 166 (1992), pp. 21–27.

[8] P. BILLINSLEY, *Probability and Measure*, 3rd ed. Wiley, NY, 1995.

[9] V.D. BLONDEL, J. THEYS, AND A.A. VLADIMIROV, *An elementary counterexample to the finiteness conjecture*, SIAM J. Matrix Anal. Appl., 24 (2003), pp. 963–970.

[10] T. BOUSCH AND J. MAIRESSE, *Asymptotic height optimization for topical IFS, Tetris heaps and the finiteness conjecture*, J. Amer. Math. Soc., 15 (2002), pp. 77–111.

[11] R.K. BRAYTON AND C.H. TONG, *Stability of dynamical systems: A constructive approach*, IEEE Trans. Circuits Syst., CAS-26 (1979), pp. 224–234.

[12] K.-L. CHUNG, *Markov Chains with Stationary Transition Probabilities*, Springer-Verlag, Berlin, 1960.

[13] A.L. COHEN, L. RODMAN, AND D.P. STANFORD, *Pointwise and uniformly convergent sets of matrices*, SIAM J. Matrix Anal. Appl., 21 (1999), pp. 93–105.

[14] X. DAI, *Extremal and Barabanov semi-norms of a semigroup generated by a bounded family of matrices*, J. Math. Anal. Appl., 379 (2011), pp. 827–833.

[15] X. DAI, *Weakly Birkhoff recurrent switching signals, almost sure and partial stability of linear switched dynamical systems*, J. Differential Equations, 250 (2011), pp. 3584–3629.

[16] X. DAI, Y. HUANG, AND M. XIAO, *Periodically switched stability induces exponential stability of discrete-time linear switched systems in the sense of Markovian probabilities*, Automatica, 47 (2011), pp. 1512–1519.

[17] X. DAI, Y. HUANG AND M. XIAO, *Realization of joint spectral radius via ergodic theory*, Electron. Res. Announc. Math. Sci., 18 (2011), 22–30.

[18] L. ELSNER, *The generalized spectral-radius theorem: An analytic-geometric proof*, Linear Algebra Appl., 220 (1995), pp. 151–159.

[19] L. GURVITS AND L. RODMAN, *Convergence of polynomially bounded semigroups of matrices*, SIAM J. Matrix Anal. Appl., 18 (1997), pp. 360–368.

[20] K.G. HARE, I.D. MORRIS, N. SIDOROV, AND J. THEYS, *An explicit counterexample to the Lagarias-Wang finiteness conjecture*, Adv. Math., 226 (2011), pp. 4667–4701.

[21] D.J. HARTFIEL, *On the limiting set of stochastic products $xA_1 \cdots A_m$*, Proc. Amer. Math. Soc., 81 (1981), pp. 201–206.

[22] D.J. HARTFIEL, *Sequential limits in Markov set-chains*, J. Appl. Probab., 28 (1991), pp. 910–913.

[23] D.J. HARTFIEL, *Nonhomogeneous Matrix Products*, World Scientific, New Jersey London, 2002.

[24] K.H. HOFMANN AND P.S. MOSTERT, *Elements of Compact Semigroups*, Charles E. Merr ll Books, Inc. Columbus, Ohio, 1966.

[25] Y. HUANG, J. LUO, T. HUANG, AND M. XIAO, *The set of stable switching sequences for discrete-time linear switched systems*, J. Math. Anal. Appl., 377 (2011), pp. 732–743.

[26] Q. HU AND D. CHENG, *Stabilizer design of planar switched linear systems*, System & Control Letters, 57 (2008), pp. 876–879.

[27] O. KALLENBERG, *Foundations of Modern Probability* (2nd Ed.), Probability and its Applications (New York). Springer-Verlag, New York, 2002.

[28] V.S. KOZYAKIN, *Algebraic unsolvability of a problem on the absolute stability of desynchronized systems*, Autom. Remote Control, 51 (1990), pp. 754–759.

[29] V.S. KOZYAKIN, *Structure of extremal trajectories of discrete linear systems and the finiteness conjecture*, Autom. Remote Control, 68 (2007), pp. 174–209.

[30] J.W. LEE AND G.E. DULLERUD, *Optimal disturbance attenuation for discrete-time switched and Markovian jump linear systems*, SIAM J. Control Optim., 45 (2006), pp. 1329–1358.

[31] J.-W. LEE AND G.E. DULLERUD, *Uniformly stabilizing sets of switching sequences for switched linear systems*, IEEE Trans. Automat. Control, **52** (2007), pp. 868–874.

[32] S. MENDENHALL AND G.L. SLATER, *A model for helicopter guidance on spiral trajectories*, in AIAA Guid. Control Conf., 1980, pp. 62–71.

[33] IAN D. MORRIS, *A rapidly-converging lower bound for the joint spectral radius via multiplicative ergodic theory*, Adv. Math., 225 (2010), pp. 3425–3445.

[34] IAN D. MORRIS, *Criterion for the stability of the finiteness property and for the uniquenss of Barabanov norms*, Linear Algebra Appl., 433 (2010), pp. 1301–1311.

[35] V.V. NEMYTSKII AND V.V. STEPANOV, *Qualitative Theory of Differential Equations*, Princeton University Press, Princeton, New Jersey 1960.

[36] M. OMLADIČ AND H. RADJAVI, *Irreducible semigroups with multiplicative spectral radius*, Linear Algebra Appl., **251** (1997), 59–72.

[37] V.I. OSELEDEC, *A multiplicative ergodic theorem, Lyapunov characteristic numbers for dynamical systems*, Trudy Mosk Mat. Obsec., 19 (1968), pp. 119–210.

[38] R. PHELPS, *Lectures on Choquet's Theorem*, Van Nostrand, Princeton, N.J., 1966.

[39] E.S. PYATNITSKIĬ AND L.B. RAPOPORT, *Periodic motion and tests for absolute stability on nonlinear nonstationary systems*, Autom. Remote Control, 52 (1991), pp. 1379–1387.

[40] G.-C. ROTA AND G. STRANG, *A note on the joint spectral radius*, Indag. Math., 22 (1960), pp. 379–381.

[41] D.P. STANFORD, *Stability for a multi-rate sampled-data system*, SIAM J. Control Optim., 17 (1979), pp. 390–399.

[42] D.P. STANFORD AND J.M. URBANO, *Some convergence properties of matrix sets*, SIAM J. Matrix Anal. Appl., 15 (1994), pp. 1132–1140.

[43] Z. SUN, *Stabilizability and insensitivity of switched linear systems*, IEEE Trans. Automat. Control, 49 (2004), pp. 1133–1137.

[44] Z. SUN, *Stabilization and optimization of switched linear systems*, Automatica, 42 (2006), pp. 783–788.

[45] Z. SUN, *A note on marginal stability of switched systems*, IEEE Trans. Automat. Control, 53 (2008), pp. 625–631.

[46] M.-H. SHIH, J.-W. WU, AND C.-T. PANG, *Asymptotic stability and generalized Gelfand spectral radius formula*, Linear Algebra Appl., 252 (1997), pp. 61–70.

[47] R. SHORTEN, F. WIRTH, O. MASON, K. WULFF, AND C. KING, *Stability criteria for switched and hybrid systems*, SIAM Rev., 49 (2007), pp. 545–592.

[48] P. WALTERS, *An Introduction to Ergodic Theory*, GTM 79, Springer-Verlag, New York, 1982.

[49] F. WIRTH, *The generalized spectral radius and extremal norms*, Linear Algebra Appl., 342 (2002), pp. 17–40.

[50] F. WIRTH, *On the structure of the extremal norms of a linear inclusion*, in: Proc. 44th IEEE Conf. Decision and Control and Eur. Control Conf., Seville, Spain, 2005, pp. 3019–3024.

[51] Z. ZHOU, *Weakly almost periodic point and measure center*, Sci. China Ser. A: Math., 36 (1992), pp. 142–53.